

# DISSERTATION

Titel

„Methodische Entwicklungen zur Qualitätssicherung  
von Programm-Evaluationen  
am Beispiel eines Lernprogramms im Hochschulbereich:  
Vienna E-Lecturing“

Verfasser

Mag. Dominik Lapka

angestrebter akademischer Grad

Doktor der Naturwissenschaften (Dr. rer. nat.)

Wien, im Mai 2010

Studienkennzahl lt. Studienblatt: A 298

Dissertationsgebiet lt. Studienblatt: Psychologie

Betreuerin: Univ.-Prof. Dr.Dr. Christiane Spiel

"Wenn Du ein Schiff bauen willst, so trommle nicht Männer zusammen, um Holz zu beschaffen, Werkzeuge vorzubereiten, Aufgaben zu vergeben und die Arbeit einzuteilen, sondern lehre die Männer die Sehnsucht nach dem weiten endlosen Meer."

Antoine de Saint-Exupéry

## **Danksagung**

Mein besonderer Dank gilt meiner Betreuerin Univ.-Prof. Dr. Dr. Christiane Spiel für Ihre Unterstützung und die zahlreichen praktischen Ratschläge, durch die mir der Einstieg in die Welt des wissenschaftlichen Publizierens erleichtert wurde.

Ich möchte mich bei meinen Kolleginnen des Vienna E-lecturing Projektes bedanken, die während der dreijährigen Projektzeit in unzähligen Gesprächen mich durch hilfreiche Anregungen und bereichernden Gedankenaustausch unterstützten. Ebenso danke ich den Mitarbeitern des Arbeitsbereiches Bildungspsychologie und Evaluation, von deren Wissen und Erfahrung ich während dieser Zeit viel profitieren konnte. Mein Dank gilt insbesondere Prof.(FH) PD Mag. Dr. Petra Wagner, die gemeinsam mit Prof. Christiane Spiel die Evaluation von Vienna E-Lecturing leitete, Ao. Univ.-Prof. Dipl.-Psych. Dr. Barbara Schober, Ass. Aus. Mag. Petra Gradingner, Mag. Dr. Dagmar Strohmeier, Dipl.-Psych. Dr. Ralph Reimann und Dr. Moira Atria.

Ein ganz besonderer Dank gilt meiner Frau Stephanie, sowie meinen Eltern, Geschwistern und Freunden, die alle auf ihre Weise zum Gelingen dieser Arbeit beigetragen haben; v.a. durch ihre Geduld und indem sie für die erforderliche Abwechslung sorgten.

---

## Inhaltsverzeichnis

<b>Danksagung .....</b>	<b>3</b>
<b>Inhaltsverzeichnis .....</b>	<b>4</b>
<b>Vorwort .....</b>	<b>6</b>
<b>1. Evaluationen im Spannungsfeld zwischen Ansprüchen der Wissenschaft und der Durchführbarkeit .....</b>	<b>7</b>
<b>2. Ziel und Aufbau der Arbeit.....</b>	<b>9</b>
<b>3. Vienna E-Lecturing.....</b>	<b>10</b>
3.1. Förderung der Lernkompetenz im akademischen Kontext.....	10
3.2. Programm-Evaluation.....	12
<b>4. Herausforderungen bei der Evaluation von VEL - Überblick der Publikationen</b>	<b>13</b>
<b>5. A Euclidean distance-based matching procedure for non-randomized comparison studies* .....</b>	<b>17</b>
5.1. Introduction.....	19
5.2. Background.....	20
5.3. The Euclidean Distance Matching Procedure, EuM.....	23
5.4. The Vienna E-Lecturing project.....	26
5.5. Results.....	29
5.6. Discussion and Outlook.....	33
5.7. References.....	36
5.7.1. Appendix: Computational steps to create distance matrices for the EuM matching procedure.....	39
<b>6. Benefits of the person-oriented perspective for program-evaluation – analyzing for differential treatment effects of the Vienna E-lecturing program* .....</b>	<b>43</b>
6.1. Introduction.....	45

---

6.2. Variable- vs. person-oriented approach.....	46
6.3. Identifying differential effects .....	48
6.4. Vienna E-Lecturing (VEL).....	50
6.5. Variable-oriented approach .....	52
6.6. Person-oriented approach .....	55
6.7. Summary and Outlook .....	62
6.8. References.....	65
<b>7. Methodenlehre: Alptraum oder Herausforderung für Psychologiestudierende?</b>	
<b>Eine Typologie auf Basis des sozialkognitiven Motivationsmodells von Dweck* .....</b>	<b>68</b>
7.1. Einleitung.....	71
7.2. Psychologiestudierende und Methodenausbildung.....	72
7.3. Alptraum oder Herausforderung – motivationspsychologische Determinanten	74
7.4. Methode.....	79
7.5. Ergebnisse .....	82
7.6. Diskussion .....	89
7.7. Literaturverzeichnis .....	96
<b>8. Conclusio .....</b>	<b>100</b>
<b>9. Literatur .....</b>	<b>104</b>
<b>Abstract.....</b>	<b>108</b>
<b>Curriculum Vitae.....</b>	<b>109</b>

---

## Vorwort

Bei der vorliegenden Arbeit handelt es sich um eine publikationsorientierte Dissertation (ÖGP, 2009) bestehend aus zwei bereits veröffentlichten Publikationen (Lapka, Wagner, Schober, Gradinger & Spiel, 2010; Spiel et al., 2008) sowie eines zur Veröffentlichung eingereichten Manuskriptes (Lapka, Wagner, Schober, Gradinger & Spiel, n.d.).

Die Publikationen sind im Rahmen des Projektes Vienna E-Lecturing\* entstanden. Vienna E-Lecturing (VEL) ist ein blended learning Programm für Psychologiestudierende, das im Rahmen einer regulären Vorlesung implementiert ist und neben Fachkompetenz auch Lernkompetenz fördert. Es wurde an der Fakultät für Psychologie der Universität Wien von Mitarbeiterinnen und Mitarbeitern des Arbeitsbereiches Bildungspsychologie und Evaluation unter der Leitung von Prof. Christiane Spiel entwickelt und erstmals im Studienjahr 2001/02 durchgeführt. Mit Hilfe der begleitenden internen Evaluation wurde das Programm laufend weiterentwickelt und optimiert (die genauen Ziele des Programms sowie der Ablauf der Evaluation sind in den folgenden Kapiteln zu finden).

Als wissenschaftlicher Projektmitarbeiter des Instituts für Wirtschaftspsychologie, Bildungspsychologie und Evaluation in den Jahren 2005 bis 2008, bestand meine Hauptaufgabe in der Evaluation spezifischer Bereiche von VEL auf der Basis der Daten der letzten Jahrgänge. Die im Rahmen dieser Tätigkeit entstandenen Analysen und Publikationen bilden die Grundlage für die vorliegende Arbeit.

\* Das Projekt wurde vom Jubiläumsfond der Österreichischen Nationalbank (Projektnummer 11107) gefördert.

---

## **1. Evaluationen im Spannungsfeld zwischen Ansprüchen der Wissenschaft und der Durchführbarkeit**

Der zunehmende wirtschaftliche Wettbewerb sowie der steigende öffentliche Druck effizient mit vorhandenen Ressourcen umzugehen führen einerseits zur Entwicklung innovativer Maßnahmen und Entstehung neuer Organisationen. Andererseits wächst auch die Legitimationspflicht gegenüber politischen Entscheidungsträgern bzw. der Öffentlichkeit (DeGEval, 2010). In diesem Zusammenhang ist immer öfter der Begriff der Evaluation zu hören. Das dabei entstehende Bewusstsein nicht alles unhinterfragt einzuführen sondern Neuentwicklungen auf Wirksamkeit und Effizienz zu prüfen ist durchaus positiv zu bewerten (Spiel, Gradinger & Lüftenegger, 2009). Leider wird der Begriff der Evaluation in diesem Zusammenhang inflationär verwendet, und es wird damit oftmals alles bezeichnet, das in irgendeiner Form mit Bewertung zu tun hat (Scriven 1991, siehe auch Spiel, Gradinger & Lüftenegger, 2009).

Dies führte in in den letzten Jahren zu einer starken Verbreitung des Begriffes der Evaluation, deren Aufgabe es ist fundierte und sachliche Informationen bereitzustellen, um eine wissensbasierte Entscheidungsfindung zu ermöglichen.

In der Praxis entsprechen jedoch eingesetzte Bewertungsverfahren oft nur eingeschränkt professionellen Evaluations-Standards und sind außerdem oft unterfinanziert (DeGEval, 2010). Fundierte und sachliche Informationen, um eine wissensbasierte Entscheidungsfindung zu ermöglichen, werden somit nur selten bereitgestellt. Doch genau das würde eine wissenschaftliche Evaluation

---

ausmachen. Denn, obwohl im Bereich der Sozialwissenschaften eine Vielzahl an unterschiedlichen Evaluations-Definitionen existiert, gibt es doch einen gemeinsamen Nenner. So spricht man dann von wissenschaftlicher Evaluation, wenn es sich um die systematische Bewertung eines Gegenstandes unter Verwendung wissenschaftlicher Methoden handelt (Westermann, 2002). Dieser wissenschaftliche Anspruch betrifft ebenso oder auch insbesondere den Bereich der Programm Evaluation (Patton, 1996; Spiel, 2005). Denn vor allem in Bereichen wie Bildung und Medizin können Ergebnisse von Programm Evaluationen zu teils gravierenden Auswirkungen führen, weshalb eine wissenschaftsbasierte Entscheidung umso wichtiger ist. Bereits seit den 70er Jahren dauert eine Auseinandersetzung von Evaluationsforschern und Sozialwissenschaftlern an, was denn nun eine effektive, wissenschaftliche Evaluation ausmacht. Daraus entwickelten sich verschiedene Richtlinien und Standards, die heute im Rahmen einer wissenschaftlichen Evaluation beachtet werden müssen (DeGEval, 2004; JCSEE, 1994; Shadish, Newman, Scheirer & Wye, 1995).

Zweifelsohne tragen wissenschaftliche Standards zu einer Qualitätssteigerung von Programm Evaluationen bei. Andererseits können sich durch die unterschiedlichen Prioritäten der Wissenschaft einerseits und der Praxis andererseits neue Spannungsfelder eröffnen. So stehen Evaluatoren\* oftmals vor der Herausforderung einerseits den wissenschaftlichen Ansprüchen zu genügen und andererseits dem vorgegebenen Kostenrahmen zu entsprechen (vgl. Spiel, Grading & Lüftenegger, 2009).

---

\* Die im Text verwendeten personenbezogenen Bezeichnungen gelten falls nicht anders angeführt für beide Geschlechter. Zwecks besserer Lesbarkeit wird nur eine Form - das generische Maskulinum - verwendet.



---

Ähnlich schwierige Herausforderungen bei der Durchsetzung wissenschaftlicher Standards gilt es bei der Evaluation von Interventionen in alltagsnahen settings zu meistern. So werden Evaluatoren in solchen natürlichen settings stärker mit der Kompliziertheit – also der Anzahl an wirkenden Faktoren – wie auch der Komplexität – also der unsicheren Wirkungsweise beteiligter Faktoren – konfrontiert (Rogers, 2008; siehe auch Scanlon & Issroff, 2005; Timmins & Miller, 2007). Dabei wirkt oftmals erschwerend, dass Evaluatoren bei der Durchsetzung wissenschaftlicher Standards in solchen Kontexten oftmals einen geringen Handlungsspielraum haben. Optimale Rahmenbedingungen sind somit nicht oder nur schwer umsetzbar, da ein zu starkes Eingreifen die Intervention an sich verändern würde (Fagan & Mihalic, 2003; Greene, Benjamin & Goodyear, 2001; Hager, Patry & Brezing, 2000; Lucke, Donald, Dower & Raphael, 2001).

Vor dem Hintergrund dieser Spannungsfelder ist folgende Arbeit entstanden. Konkret tauchten die angesprochenen Herausforderungen in ähnlicher Form bei der Evaluation des internetgestützten Hochschulprogramms Vienna E-Lecturing auf. Um die wissenschaftlichen Standards in diesem Spannungsfeld durchsetzen zu können, machte sich das Evaluationsteam innovative Ansätze zu Nutze, welche im Rahmen dieser Arbeit vorgestellt werden und auch bei Evaluationen in anderen Bereichen angewendet werden können.

## **2. Ziel und Aufbau der Arbeit**

Ziel der vorliegenden Dissertation ist somit, methodische Ansätze zur Qualitätssicherung von Programm Evaluationen vorzustellen. Dabei werden zwei konkrete Maßnahmen aufgezeigt, die zur höheren Qualität von Programm

---

Evaluationen beitragen und gleichzeitig das Spannungsfeld zwischen Wissenschaftlichkeit und Durchführbarkeit entladen. Dazu zählen:

1. Ein neues, einfach durchführbares Verfahren zum Matchen von Versuchs- und Kontrollpersonen bei nicht durchführbarer Randomisierung und
2. wird im Rahmen dieser Dissertation für den ergänzenden Einsatz personenorientierter Analysen plädiert. Diese Ergänzung zu den üblichen variablenorientierten Analysen prüft eine mögliche differentielle Wirksamkeit eines Programms.

Um beide Maßnahmen in der Anwendung auf ihre Nützlichkeit zu prüfen dient als Anschauungsbeispiel die Evaluation des Hochschulprogramms Vienna E-Lecturing. Die Ziele, der Ablauf sowie der theoretische Hintergrund des Programms werden im Folgenden kurz beschrieben. Anschließend folgt ein Überblick über die begleitende Programm-Evaluation. Die dabei auftretenden Herausforderungen, die zur Implementierung der vorgestellten Ansätze führten, werden in einer einleitenden Zusammenfassung der drei Publikationen kurz umrissen. Die schrittweise Umsetzung der entwickelten Lösungsansätze, sowie genaue Studienergebnisse sind in den jeweiligen Publikationen enthalten. Abschließend werden in einer zusammenfassenden Conclusio Parallelen zur Literatur gezogen und weiterführende Forschungsfragen gestellt.

### **3. Vienna E-Lecturing**

#### **3.1. Förderung der Lernkompetenz im akademischen Kontext**

Als Anschauungsbeispiel für die vorgestellten Methoden zur Qualitätssicherung von Programm-Evaluationen dient die Evaluation von „Vienna

---

E-Lecturing" (VEL), einem internetgestützten Lernprogramm für Psychologiestudierende, das an der Universität Wien durchgeführt wird. Das Programm ist in eine laufende zweistündige Pflichtlehrveranstaltung des zweiten Studienabschnittes (Hauptstudium) eingebettet, die sich mit den Inhalten von „Forschungsmethoden und Evaluation" befasst und sich über ein gesamtes Studienjahr (Winter- und Sommersemester) erstreckt. Speziell diese Domäne ist für die Realisierung neuer Lehr-Lernformen besonders bedeutsam, da die Mehrheit der Psychologiestudierenden der verpflichtenden Methodenausbildung im Rahmen ihres Studiums negativ gegenüber stehen (Gal & Ginsburg, 1994). Man kann davon ausgehen, dass 70 bis 80 Prozent der Studierenden eine für den Lernerfolg hinderliche Statistikangst erleben (Onwuegbuzie, 2004; Zeidner, 1991). Vor diesem Hintergrund entstand das VEL Programm, das in mehreren Phasen entwickelt, evaluiert und optimiert wurde (Spiel et al., 2004).

In VEL wurde über die Vermittlung des reinen Fachwissens hinausgehend das Lernen als Solches zum expliziten Förderungsgegenstand gemacht. Konkret verfolgt VEL vier Lehrziele: Vermittlung von (1) Fachwissen, (2) Lernkompetenz, (3) Kompetenz zum kollaborativen Lernen und (4) E-Kompetenz. Deshalb wurde die Gestaltung der Lehrveranstaltung nach instruktions- und motivationspsychologischen Gesichtspunkten weitmöglichst optimiert und das gemeinsame Arbeiten in Lerngruppen zu einem grundlegenden Prinzip erhoben. E-Learning als eine für viele Studierende ansprechende Vermittlungsform sollte dazu beitragen, fachbezogene Berührungsängste und motivationale Defizite zu reduzieren. Dabei wurde geachtet Online-Module mit realen Präsenzeinheiten systematisch miteinander zu kombinieren. In der Gesamtheit, also im Verlauf eines Studienjahres (Winter- und Sommersemester), umfasst VEL 10 Online-Module, welche Literatur, inhaltliche Aufgaben, Selbstregulationsübungen und

---

Selbsttests mit einem individuellen Feedback beinhalten sowie zusätzlich 15 Präsenzeinheiten, aufgegliedert in Vorlesungen, Seminare und Tutorien. Das genaue Trainingskonzept wurde von Wagner, Schober, Reimann, Atria und Spiel (2007) publiziert.

### 3.2. Programm-Evaluation

Die komplexe Begleitforschung zum Vienna E-Lecturing Programm umfasst unter anderem eine summative Evaluation, welche die Wirksamkeit von VEL bezüglich der Zielerreichung prüft - der Schwerpunkt liegt dabei auf dem Fachwissen und der Lernkompetenz. Für die summative Evaluation wurde ein 2-Gruppen-Meßwiederholungsdesign eingesetzt. Die Vergleichsgruppe nahm an einer reduzierten Variante von VEL teil. Diese entspricht im Aufbau einer regulären Vorlesung mit zusätzlicher geringer Online-Unterstützung, die einen Selbsttest sowie die notwendige Literatur beinhaltet. In dieser Version finden keine Seminare und Tutorien statt und ebenso ist der Bereich der Gruppenaufgaben und Selbstregulationsförderung ausgeklammert. Die Zuteilung zu einer der beiden Gruppen erfolgt freiwillig über eine schriftliche Bewerbung und ein Motivationsschreiben.

Die Erhebungen erfolgten Anfang des Wintersemesters zu Beginn der Lehrveranstaltung (Pre-Test), am Ende des Wintersemesters (Post-Test 1) und am Ende des Sommersemesters (Post-Test 2). Dabei wurde sowohl die Fachkompetenz als auch die Lernkompetenz erhoben. Die Fachkompetenz wurde beim Pre-Test mittels eines eigens konstruierten Vorwissenstests erfasst, zu den beiden Post-Tests wurden die Leistungen der jeweiligen Vorlesungsprüfung herangezogen. Zur Erfassung der Lernkompetenz wurde ein Inventar aus

---

verschiedenen publizierten Skalen erstellt, das Lernkompetenz auf den Ebenen der Metakognition, Kognition, Motivation und Emotion erfasst.

Das Vorgehen erfolgte gemäß der wissenschaftlichen Kriterien einer Programmevaluation (siehe z. B. Hager, Patry & Brezing, 2000; Rossi, Freeman & Lipsey, 1999; Spiel, 2005). Dabei ergaben sich aufgrund der Alltagsnähe des Lernprogramms einige Schwierigkeiten - beispielsweise war die randomisierte Zuteilung der Studierenden zu Versuchs- oder Vergleichsgruppe nicht durchführbar. Die konkreten Herausforderungen werden in der folgenden Zusammenfassung der drei Publikationen kurz umrissen.

#### **4. Herausforderungen bei der Evaluation von VEL - Überblick der Publikationen**

Vom wissenschaftlichen Standpunkt aus ist die Randomisierung das Mittel der Wahl, um personengebundene Störvariablen zu kontrollieren, was wiederum für die Interpretation der beobachteten Effekte als Wirkungen der evaluierten Maßnahme notwendig ist (Bortz & Döring, 2002). Andernfalls wären allfällige Unterschiede durch Selbstselektionseffekte zu erklären.

Experimentelle Designs sind jedoch im Bereich der Sozialpolitik eher die Ausnahme (Stegmann, 2009). Ebenso ist bei Interventionen in natürlichen settings eine randomisierung Zuteilung oft nicht (oder nur schwer) möglich. Auf Grund der freiwilligen Bewerbung der Studierenden zum VEL Programm war die randomisierte Zuteilung der Studierenden zu Versuchs- oder Vergleichsgruppe auch hier nicht möglich. Wenn ein experimentelles Design nicht realisierbar ist und somit Eigenschaften, welche die Wirkung der Intervention beeinflussen können nicht zufällig auf Versuchs- und Kontrollgruppe verteilt sind, werden in

---

der Literatur einige Alternativen genannt, um trotzdem eine kausale Wirkung des Programms feststellen zu können. Beispiele dafür sind das statistisches matching (Rubin, 1977) oder propensity score matching (D'Agostino, 2005). Beide Methoden haben jedoch ihre Einschränkungen. So ist das statistische matching nur bei wenigen zu kontrollierenden Variablen möglich (Dehejia & Wahba, 2002) und die Berechnung von propensity scores kann sehr komplex werden. Im Rahmen des Projektes VEL wurde eine einfache Alternative entwickelt, die mehrere intervallskalierte Variablen gleichzeitig berücksichtigen kann: das Euclidean Distance Matching (EuM). Die dahinter stehende Idee ist ähnlich der der meisten anderen matching Verfahren: gleiche oder sehr ähnliche Personenpaare zu finden, um eine randomisierte Zuteilung zu den beiden Gruppen zu simulieren. Unterschiede zwischen Versuchs- und Vergleichsgruppe können in weiterer Folge als Folge der Intervention interpretiert werden. In der Publikation wird die schrittweise Durchführung dieser Prozedur genau erklärt und am Beispiel von VEL exemplarisch dargestellt. Ein vorher-nachher Vergleich zeigt, dass nach der Durchführung der matching Prozedur sich die Versuchs-Gruppe von ihrer gematchten Vergleichs-Gruppe zu Beginn der Intervention in keiner der relevanten Variablen unterscheidet. Die EuM-Prozedur wird als wirksames Instrument vorgestellt, für das keine fortgeschrittenen statistischen Kenntnisse notwendig sind.

Eine andere Herausforderung bei der Evaluation von Interventionen im alltagsnahen Kontext und e-learning Umgebungen und somit auch bei VEL, liegt in der Komplexität der zusammenspielenden Faktoren (Scanlon & Issroff, 2005). Meist sind in solchen settings klare und starke Interventionseffekte nur schwer zu finden. Die Ursachen dafür können vielschichtig sein: das Training ist nicht bzw. nur selektiv wirksam, die Evaluation entspricht nicht den etablierten

---

Standards, die Erhebungsinstrumente sind nicht änderungssensitiv, etc. (Bierman, 2006; JCSEE, 1994). So liegt der zweite Fokus der vorliegenden Arbeit bei einem der möglichen Gründe, nämlich bei der differentiellen Wirksamkeit von Interventionsmaßnahmen und deren Prüfung durch ergänzende personenorientierte Analysen, wie sie auch bei der Evaluation von VEL durchgeführt wurden. Konkret plädiert der zweite Beitrag für den Einsatz des personenorientierten Zugangs in Ergänzung zum gebräuchlichen variablenorientierten Zugang der Datenanalyse, zur Identifikation differentieller Interventionseffekte. Der personenorientierte Ansatz wurde in den letzten Jahren vor allem von der Arbeitsgruppe um Magnusson und Bergman proklamiert (z.B. Bergman & Magnusson, 1997; Magnusson, 2000; siehe auch von Eye, Bogat & Rhodes, 2006) und basiert auf der Annahme, dass in der heterogenen Interventionsgruppe homogene Subgruppen existieren, und dass aggregierte Werte für die gesamte Gruppe dem differentiellen Bild auf Subgruppen-Ebene widerspricht (von Eye, 2006). Sowohl die Anwendung des variablen-orientierten als auch des personen-orientierten Zugangs, werden an Hand des Illustrationsbeispiels VEL vorgestellt, die jeweiligen Ergebnisse miteinander verglichen und die Vorteile aufgezeigt.

Der dritte Beitrag leistet Vorarbeit für die personenorientierten Analysen indem unter der Anwendung clusteranalytischer Verfahren unterschiedliche Subgruppen von Studierenden identifiziert werden. Um einerseits artifizielle Typen zu vermeiden (Bortz, 1999) und andererseits Subgruppen zu identifizieren, auf die VEL eine differentielle Wirkung erzielen könnte, ist der Bezug zur Theorie essentiell. Durch die oftmals problematische motivationale Situation von Psychologie-Studierenden im Bereich Statistik und Forschungsmethoden (Birenbaum & Eylath, 1994; Onwuegbuzie, 2004;

---

Onwuegbuzie & Wilson, 2003; Zeidner, 1991) ist der Bezug zu einem Motivationsmodell naheliegend. Konkret wird als theoretische Grundlage für die Variablenauswahl und Interpretation der Ergebnisse der Clusteranalyse auf das sozialkognitive Motivationsmodell (Dweck & Leggett, 1988) Bezug genommen, auf dem auch die Motivationsförderung bei VEL basiert. Neben der Identifizierung von Subgruppen liegen weitere Ziele dieser Publikation (1) in der Beschreibung der motivationalen Situation Psychologiestudierender im Fach Methodenlehre und (2) in der Prüfung der zentralen Annahmen des sozialkognitiven Motivationsmodells (Dweck & Leggett, 1988). Die Ergebnisse legen u.a. den Schluss nahe, dass ein Lernprogramm wie VEL auf die drei resultierenden Subgruppen unterschiedlich wirken kann. Des Weiteren werden im Rahmen dieser Publikation Förderansätze abgeleitet sowie eine Erweiterung des Modells empfohlen.

In den folgenden Publikationen, ist die schrittweise Umsetzung der jeweiligen Maßnahmen dargestellt, so dass sie auch als Anleitung für Evaluatoren dienen können. Des Weiteren werden dort die genauen Studienergebnisse präsentiert, welche u.a. die Frage beantworten, was die Maßnahmen in Bezug auf die Evaluation von VEL an Optimierung einbrachten. Daran anschließend werden in einer zusammenfassenden Conclusio Parallelen zur aktuellen Literatur gezogen und weiterführende Forschungsfragen genannt.



---

## **5. A Euclidean distance-based matching procedure for non-randomized comparison studies\***

Spiel, C., Lapka, D., Gradinger, P., Zodlhofer, E. M., Reimann, R., Schober, B., Wagner, P.<sup>1</sup>, & von Eye, A.<sup>2</sup> (2008). A euclidean distance-based matching procedure for non-randomized comparison studies. *European Psychologist*, 13 (3), 180-187.

University of Vienna

<sup>1</sup> University of Applied Sciences Linz

<sup>2</sup> Michigan State University

\* This research was supported in part by the "Jubilee Fund of the Austrian National Bank", Grant No. 11107 and by the Austrian Federal Ministry for Education, Science and Culture. Address correspondence concerning this article to Christiane Spiel, University of Vienna, Faculty of Psychology, Universitaetsstr. 7, A-1010 Vienna, Austria, phone: +43 1 4277 47870, fax: +43 1 4277 47879.

---

## Abstract

For intervention programs that are applied in natural settings, randomization often is difficult or impossible to achieve. If treated individuals are compared with individuals from a non-randomized comparison group, the inference of causality can be biased. Similar distributions in the relevant characteristics of the treatment and the comparison groups cannot be expected. To adjust between-group comparisons for pre-existing differences, this article proposes a simple matching procedure. This procedure involves pairing of treatment and comparison individuals based on observable characteristics, using Euclidean distance scores. Application of the proposed Euclidean-distance Matching (EuM) procedure to data from the Viennese E-Lecturing (VEL) project yields satisfying results. Possible generalizations and applications of the EuM procedure are discussed.

**Keywords:** Matching, randomization, nonexperimental design, Euclidean distance, e-learning.

---

## 5.1. Introduction

Many intervention programs are applied in natural settings. In each application, even without extensive experimental control, causal treatment effects are expected. In evaluation studies, conducted to demonstrate program effectiveness, randomized assignment of individuals to treatment and control groups, to dosage groups, or to types of treatment is often impossible. In many evaluation studies, a treatment group is compared to a nonexperimental, that is, non-randomized comparison group, the control group. Although the term control group typically describes an experimental comparison group, it, is also frequently used in nonexperimental (non-randomized) designs. The inference of causality in studies of this type can be biased. One can expect that treatment effects are confounded with factors that determine self-selection into treatment or control groups. In this article, we propose a simple matching procedure. This procedure allows researchers to pair treatment and control individuals based on information that exists prior to treatment. The matches are created using Euclidean distance scores.

This article is organized as follows. In the first section, we discuss prior research on adjusting between-group comparisons for pre-existing differences. In the second section, the Euclidean Distance Matching procedure (EuM) is presented. The EuM procedure was developed within the Vienna E-Lecturing (VEL) project which is described in Section three. In Section four, we use data from this project to illustrate our matching procedure. In Section five, we discuss results, generalizations, and possible applications of the EuM procedure.

---

## 5.2. Background

To demonstrate effectiveness of intervention programs, outcome evaluation is often performed (Berk & Rossi, 1999; Fink, 1995; Spiel, 2005). Accuracy standards as defined by the Joint Committee on Standards for Educational Evaluation (JCSEE, 1994; see DeGEval, 2002) are intended to ensure that an evaluation reveals and transmits accurate information about a program's merits. Based on these standards, it is determined whether an evaluation has produced sound information.

Intervention programs are expected to have causal effects. Holland (1986) views a cause as a manipulation or treatment that brings about change in the variable of interest, compared to some baseline (cf. Cox, 1992). The variable of interest is observed under treatment and under control conditions. To prevent bias, sound outcome evaluation performs randomized assignment of individuals to treatment or control conditions, or to different levels or types of treatment. However, for many intervention programs, in natural settings, randomization is difficult or impossible to achieve. Therefore, often, only a treatment pre-post test design is applied or treated individuals are compared with individuals from a nonexperimental comparison group. In addition, self-selection takes place. One problem with this procedure is that self-selection into treatment or control groups, or the researchers' decisions when assigning individuals to treatment and control groups can cause bias. Similar distributions in the relevant characteristics of the treatment and the comparison groups cannot be expected. Thus, comparisons may confound the effect of the treatment with the effect of the factors that guided selection and group assignment.

---

A number of procedures has been proposed to adjust between-group comparisons for pre-existing differences. One common approach involves application of standard linear regression. Typically, the various predictors of outcomes are included as regressors (covariates). However, the regression approach has its limitations (Heckman, Ichimura & Todd, 1998; Rosenbaum & Rubin, 1983). Specifically, the comparison groups may not only differ in the means but also in the distributions of the covariates, and create complications such as heterogeneous regression slopes and non-normal residual distributions. If this is the case, regression adjustment may not work as intended.

An alternative approach to regression involves statistical matching. To substitute for the absence of a randomized experimental control group, data are used from individuals with the same (or very similar) scores on observed pre-treatment covariates as the treated individuals. Comparing two individuals, one treated and one not treated, with the same observed characteristics, is assumed to mimic the comparison of these two individuals in a randomized experiment (Rubin, 1977). Matching individuals based on their vector of covariates requires a sufficiently large comparison group. Matching with a small number of characteristics is straightforward. However, the process can become quite involved if matching is performed based on many characteristics. The difficulty of finding exact or close matches for each of the treated individuals increases with the number of covariates (Dehejia & Wahba, 2002).

To address this issue, the use of the *propensity scores* has been propagated (D'Agostino, 1998, 2005; Rosenbaum & Rubin, 1983). The propensity score of an individual is defined as the conditional probability of being treated, given this individual's covariate scores. More specifically and assuming complete data, the

---

propensity score for individual  $i$ , for  $i = 1, \dots, N$  is the conditional probability that  $i$  was assigned to a particular treatment,  $Z_i$ , with  $Z_i = \{0, 1\}$ , and 1 indicating the treatment condition. The probability is conditioned on the individual's vector of covariates,  $x_i$ , or

$$e(x_i) = pr(Z_i = 1 | X_i = x_i).$$

Given the  $X$  scores, the  $Z_i$  are independent, or

$$pr(Z_1 = z_1, \dots, Z_n = z_n | X_1 = x_1, \dots, X_n = x_n) = \prod_{i=1}^N e((x_i)^{z_i} (1 - e(x_i))^{1-z_i}).$$

D'Agostino (2005) notes that, for a specific value of the propensity score, the difference between the treatment and the control means for all individuals with this propensity score represents an unbiased estimate of the average treatment effect.

Propensity scores can be estimated using discriminant analysis or logistic regression. For discriminant analysis, the observed covariates must be assumed to be multivariate normal, conditional on  $Z$ . For logistic regression, this assumption is not necessary. Currently, propensity scores are used in the context of matching, stratification, and regression adjustment (for more detail, see Dehejia & Wahba, 2002; Foster, 2003; Heckman, Ichimura, & Todd, 1998). However, although propensity scores can be estimated using most general purpose statistical software packages, diagnostics is not routine and can be complex. This applies in particular to methods that examine the sensitivity of the estimated treatment effect to small changes in the propensity score specification (see Dehejia, 2005).

Propensity scores were introduced for a number of reasons, one of the most important being that matching based on the raw scores of multiple covariates is

---

complicated. In this article, we propose a simple procedure that allows one to perform matching based on multiple raw covariate scores. The procedure neither has the limitations of the regression approach nor needs statistical expertise as advanced as the application of propensity scores. The matches are created using Euclidean distance scores. The steps of this procedure can also be performed using standard statistical software.

### 5.3. The Euclidean Distance Matching Procedure, EuM

For proper application of the EuM procedure, two conditions must be fulfilled. The proposed EuM procedure is (1) based on the assumption that the changes that intervention programs are expected to cause be measured at an interval or ratio scale level. Categorical variables can be considered after dummy coding. For the EuM procedure, it is not required that scores be normally distributed. The EuM procedure (2) requires complete data. To assure complete data, it is recommended to estimate and impute missing data using methods of multiple imputation (Little & Rubin, 2002; Schafer & Graham, 2002). Alternatively, one-step maximum likelihood estimation and the hot deck method are recommended (Little & Rubin, 2002). Estimation and imputation are more complex when missingness is non-random. One can consider performing estimation and imputation separately for treatment and control groups (or any other strata) when missingness varies with group membership. Eliminating individuals casewise or listwise, or mean imputation are “editorial methods” that are known to do more harm than good.

The idea that carries the EuM procedure is comparable to the idea that carries propensity score matching and most other matching procedures. Finding

---

exact or very close matches creates a situation in which assignment to treatment and control groups can be considered random. Differences between treatment and control groups can then be interpreted as caused by the treatment. The following paragraphs provide a description of the EuM procedure in consecutive steps. The procedure involves four steps.

*Step 1: Variable selection.* The first step involves the selection of the variables used for matching. These variables must be scaled, at the least, at the interval level, and they must be comparable in their scaling. The selection of variables is based on theory. Specifically, these variables are selected based on their theoretical relationships with the outcome variables. Matching based on different variable selections can lead to different results. Note that this argument applies accordingly to the selection of variables on which estimation and imputation of missing values are based (see Step 2). To avoid possible biases, it is important to check for multicollinearity among the selected variables. As method, a multiple regression in a standard package such as SPSS or SYSTAT could be applied and collinearity diagnostics – in particular the tolerance values ( $1 - R^2$ ) – used for checking. Ideally, all tolerance values should be high (certainly more than 0.2 and preferably much closer to 1). If the tolerance value of a selected variable is low, exclusion of this variable should be considered. Alternatively, orthogonalization using principal components analysis of the selected variables is an option.

*Step 2: Data preparation.* In the second step, data are cleaned and completed. In most cases, completion involves estimation and imputation of missing values. In addition, because matching is based on the Euclidean distance between individuals in the multivariate data space, the variables should be



---

placed on commensurable scales to prevent variables with larger variances and variables scored using larger numbers from dominating the calculated distance. Standardization is a very simple and often used method to create commensurable scales. It is important that the standardization use the entire sample. If standardization is performed separately for the strata, differences in means and standard deviations as well as the effects of these differences on the outcomes may be obscured. Percentage of maximum possible (POMP) is an alternative method, that scales all variables between 0 and 100 (Cohen, Cohen, Aiken, & West, 1999).

*Step 3: Calculation of Euclidean distances.* Step 3 involves the calculation of a distance matrix. This matrix is of size  $N \times N$ . Consider the two points,  $P = (p_1, p_2, \dots, p_d)$  and  $Q = (q_1, q_2, \dots, q_d)$ , where  $d$  indicates the number of variables used to calculate the distance. Then, the  $d$ -dimensional Euclidean distance between  $P$  and  $Q$  is

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_d - q_d)^2} = \sqrt{\sum_{i=1}^d (p_i - q_i)^2}$$

The distance matrix contains the between-individual distances in the space of those  $d$  variables on which the matching is based.

*Step 4: Matching.* This step uses the classification of individuals as members of the treatment and comparison groups. For each member of the treatment group, possible matches are identified. In the ideal case, there is more than one match for each individual in the treatment group. In one extreme case, the distance between two individuals is zero. In this case, we have a perfect match. In the other extreme case, the distance is large. Whenever there is no zero-distance match for a member of the treatment group, researchers will have to

---

decide whether a match for this individual exists. The result of the matching procedure is a data set that contains a variable that identifies the matches. In small samples, the probability is high that not every individual will have a match.

In the literature, a number of proposals has been made concerning cut offs. For example, in *caliper matching* (Cochran & Rubin, 1973; for an application, see Chabris & Glickman, 2006), a strict cut off of 0.15 of a standard deviation is often used (see also *kernel-based matching*; Heckman, Ichimura, & Todd, 1998). Cut offs are easily given, but they always carry an arbitrary element. For example, the probability of exact matches and low distance scores decreases as the number of matching variables increases. The desiderate of close or perfect matches is invoked if matched pairs comparisons are performed. If matched groups comparisons are performed, this desiderate is shifted to the level of means and standard deviations. In both cases, mean differences in the matching variables between the treatment and the matched control groups clearly indicate unsuccessful matching.

In the next section, we present an application of the EuM procedure to data from the the Vienna E-Lecturing (VEL) project. The VEL is described in some detail to illustrate why randomization was impossible to achieve and to make selection of the matching variables plausible.

#### 5.4. The Vienna E-Lecturing project

It is well known and has repeatedly been shown that methodology and statistics courses are not among the most popular with students majoring in Psychology. In addition, these courses are often associated with the highest

---

levels of anxiety (Blalock, 1987; Gal & Ginsburg, 1994; Onwuegbuzie, 2000). It has been assumed that this state of affairs results in negative consequences for competence development and academic progress as indicated by poor performance in these courses. In addition, these attitudes may have effects on the use of statistics and methodology by these students (Birenbaum & Eylath, 1994; Onwuegbuzie, 2000).

To overcome the discrepancy between relevance and acceptance of the methodology, the Internet-based teaching concept Vienna E-Learning (VEL) was developed, evaluated, and optimized over several phases (Spiel et al., 2004). The VEL was implemented in the mandatory two-semester course on research methods and evaluation that is offered for juniors majoring in psychology at the University of Vienna.

The teaching concept of the VEL is based on principles developed in motivational and instructional psychology (Pintrich & Schunk, 1996). In addition to imparting factual knowledge in the area of methodology, the VEL aims at mediating learning competence, collaborative learning abilities, and competence in computing (= e-competence). The didactical principles that are implemented to achieve these goals involve networking, focusing on the various learning goals, on collaborative learning in groups of five students who work on specific tasks, and the on-site discussion of online modules with instructors.

There is a total of 10 online modules. These modules are implemented on the WebCT Vista learning platform that is supported by the computing center of the university. The modules cover the content of the Research Methods and Evaluation course, and are presented in a uniform design. The tasks are

---

presented within the modules (for more detail, see Schober, Wagner, Reimann, Atria, & Spiel, 2006).

The on-site meetings include a kick off event, six “meet the expert” units, four tutorials, and four self-regulated learning units. The meet the expert units serve to discuss content at a deeper level. In the tutorials, the fundamentals of group work are discussed, learning groups are formed, basic computing skills are taught, group work is discussed, and project management skills relevant to performing evaluations are taught. The self-regulated learning units provide concrete guidance on the self-monitoring of learning. Based on Schmitz’ (2001) process model of self-regulated learning, self-regulation, self-motivation, dealing with feedback, emotion regulation, goal setting, and learning strategies are trained (for more detail, see Wagner, Schober, Reimann, Atria, & Spiel, in press).

After being developed in a number of research projects, the VEL concept was integrated in the required course on research methods and evaluation. It is now a component of the standard curriculum at the University of Vienna. The structure of the VEL is based on learning theory. It is dynamic in the sense that it can be adapted and optimized in regard to both content and student progress. The evaluation of the VEL focuses on effectiveness. The data used in this article were collected in the context of a comprehensive evaluation of a research project that involved both formative and summative evaluation.

The realization of the extensive learning program of the VEL is cost-intensive. Therefore, it can be offered to no more than 50 students per year, that is, 10 learning groups. The remaining students are presented a reduced version of the VEL. This version does not contain guided group work and self-regulated learning training. Randomized assignment to the two versions of the VEL is impossible.

---

Students have to submit an application to participate in the context of the full version of the VEL. Therefore, the estimation of the effectiveness of the VEL can be biased by self-selection.

Thus far, data from two cohorts ( $n = 90$ ; 15 males, 75 females) of participants in the full version of the VEL are available in addition to data from 375 students (50 males, 325 females) who participated using the reduced version. The summative evaluation of the effectiveness of the VEL involves one pre-test and two post-tests, the latter two conducted after the first and the second semesters. Data were collected with focus on factual knowledge. In addition, variables describing motivation, self-regulated learning competence, collaborative learning and e-competence, as well as changes in these variables were observed. To take into account possible pre-existing differences, the Euclidean distance Matching (EuM) procedure was used. In the next section, we present results of the application of EuM.

## 5.5. Results

In Step 1, the variables relevant for matching were determined. Based on theory, results reported in the literature, and the goals of the VEL, the following five variables were selected: factual knowledge in statistics, research methods, and evaluation, interest in the topic of the course, monitoring the learning process, organization of the self-regulated learning process, and time management. The factual knowledge score ranges from 0 to 27 points. All other variables were scaled on Likert-type scales ranging from 1 = completely disagree to 6 = completely agree. Table 1 displays, for each of the two groups (intensive and standard), the means and standard deviations of these variables before

matching, the mean differences between the two groups with the 95% confidence intervals, and t-tests comparing the two groups. Group means of all variables differ from one another more than 0.15 of the respective standard deviations (Cochran & Rubin, 1973). In addition, to check for multicollinearity a multiple regression was applied and the tolerance values of the five selected variables were computed (see Table 2). All tolerance values are high and close to 1. The lowest value is .839. We, therefore, conclude that multicollinearity is not an issue of major concern and we keep all five variables in the matching procedure.

*Table 1: Mean level comparison of the intensive-group and the standard-group before the matching procedure*

	intensive (n=90)		standard (n=375)		t-test			mean differ- ence	95% CI	
Scales	M	SD	M	SD	<i>t</i>	<i>df</i>	<i>p</i>			
Factual knowledge	15.33	3.50	14.24	3.52	2.66	463	.008*	1.10	0.29	1.91
Interest	3.96	1.05	3.64	0.96	2.76	463	.006*	0.32	0.09	0.54
Monitoring	3.38	0.85	3.23	0.89	1.45	463	.148	0.15	-0.05	0.35
Organization	4.03	1.29	3.83	1.30	1.27	463	.206	0.19	-0.11	0.49
Time management	3.60	1.14	3.91	1.15	-2.28	463	.023*	-0.31	-0.57	-0.04

*Note:* The table contains mean-scores (M) and standard deviations (SD) of the two groups, the results of the two-tailed t-tests (*t*, *df*, *p*) as well as the mean differences with the 95% confidence interval. The scales (except factual knowledge in statistics) range from 1 (= completely disagree) to 6 (= completely agree). The factual knowledge score ranges from 0 to 27 points. Significant mean differences ( $p < 0.05$ ) are indicated by an asterisk.

In Step 2, the students who only participated in the first wave of data collection were eliminated. The remaining missing values were estimated and imputed using the EM algorithm available in SYSTAT 12.0 (2007). From the intensive group, 87 out of 90 students remained in the data set. From the standard-group, 306 of the 375 participated in at least two waves of the data

collection. The scales of the variables used for matching were standardized. Then, in Step 3, the Euclidean Distances were calculated to create a 393 x 393 distance matrix. Computational steps are given in the appendix. In Step 4, the matches were identified separately for men and women (based on previous studies, we expected gender differences). Matches were identified for 81 individuals (intensive group: 12 males, 69 females; standard group: 12 males, 69 females). Each subject in the intensive group was matched with a subject in the standard group having the lowest Euclidean Distance score (the distance score matrix can be requested from the first author).

*Table 2: Multiple Regression with Factual knowledge, Interest, Monitoring, Organization and Time management.*

Scales	R <sup>2</sup>	Tolerance value (1- R <sup>2</sup> )
Factual knowledge	.022	.978
Interest	.161	.839
Monitoring	.079	.921
Organization	.126	.874
Time management	.061	.939

*Note:* The table contains the R<sup>2</sup> and the tolerance values of the regression analysis for each variable with the others as predictors.

The result of the matching procedure is a data set that contains a variable that identifies the matches. Because matching involved five variables, we did not expect very close matches (matching scores:  $M = .98$ ,  $SD = .31$ ). Therefore, for the summative evaluation of the effectiveness of VEL, we applied matched groups comparisons instead of matched pairs comparisons. The application of the EuM generated a matched standard-group with no significant differences in the relevant variables to the intensive-group. Moreover, the group means of all variables differ less than 0.15 of the respective standard deviations (see Cochran

& Rubin, 1973). In addition, no differences in the standard deviations were found. Table 3 shows descriptive statistics for both groups after matching.

*Table 3: Mean level comparison of the intensive-group and the standard-group after the matching procedure*

	Intensive (n=81)		Standard (n=81)		Levene-test		t-test			mean differ- ence	95% CI	
Scales	M	SD	M	SD	F	p	t	df	p			
Factual knowledge	15.36	3.16	15.20	2.97	.117	.73	.333	160	.74	0.16	-0.79	1.12
Interest	3.92	1.00	3.88	.88	2.078	.15	.257	160	.80	0.04	-0.25	0.33
Monitoring	3.35	.80	3.28	.76	.005	.95	.557	160	.58	0.07	-0.17	0.31
Organization	3.95	1.27	3.98	1.18	1.447	.23	-.128	160	.90	-0.02	-0.40	0.36
Time management	3.62	1.17	3.78	1.09	.030	.86	-.891	160	.37	-0.16	-0.51	0.19

*Note:* The table contains means (M) and standard deviations (SD) of the two groups, the results of Levene-test ( $F$ ,  $p$ ) and the two-tailed t-tests ( $t$ ,  $df$ ,  $p$ ), as well as the mean differences with the 95% confidence interval. The scales (except factual knowledge) range from 1 (= completely disagree) to 6 (= completely agree). The factual knowledge score ranges from 0 to 27 points.

The results in Table 3 show that the results in Table 1 come with considerable bias. In different words, without matching, the participants of extended learning program seem to outperform the other students in a number of variables, before the start of the learning program. The matching procedure allowed us to create two groups that are equivalent on the five matching variables.

Sample results of the summative evaluation of the VEL showed higher factual knowledge in the intensive group ( $M=76.11$ ;  $SD=10.18$ ) than in the control group ( $M=72.30$ ;  $SD=10.94$ ),  $t(648) = 3.032$ ,  $p < .05$ ; Mean difference = 3.81 (CI 1.34, 6.27) for the non-matched groups comparison after the first term, but also for the matched groups comparison (intensive group:  $M=76.31$ ,  $SD=9.94$ ;



---

control group:  $M=72.92$ ,  $SD=10.19$ ;  $t(154) = 2.086$ ;  $p < .05$ ); Mean difference = 3.38 (CI 0.18, 6.59). Clearly, these results look similar. However, after matching, the mean differences are no longer confounded with a priori existing differences in the matching variables.

The last section discusses advantages and limitations of the proposed Euclidean-distance matching procedure, and proposes generalisations and applications of the EuM procedure.

## 5.6. Discussion and Outlook

For many psychological intervention programs, randomization is difficult or impossible to achieve. If this is the case, it is difficult to infer causality just by way of comparing a treatment group to a nonexperimental (non-randomized) comparison group. In particular, it can be expected that self-selection of the individuals into one of the conditions causes bias. To correct for sample selection bias due to a priori differences between the treatment and the comparison groups, matching procedures are used. The procedure presented in this paper matches each individual in the treatment group with one individual in the comparison group using multidimensional Euclidean distance scores.

Application of the proposed EuM procedure to data from the Vienna E-Lecturing (VEL) project shows clear results. The intensive-program group and the matched standard group, both self-selected, do not differ in the baseline in any of the relevant variables used for matching. Furthermore, the cut off criterion of 0.15 of a standard deviation (Cochran & Rubin, 1973) is reached for the group means of all matching variables. Sample results of the summative evaluation of

---

the VEL showed not only for the non-matched groups comparison higher factual knowledge in the intensive group after the first term, but also for the matched groups comparison. That means that conclusions are similar after potential confounding is removed.

The new EuM procedure does not require advanced statistical expertise. This is seen as an advantage of the EuM procedure, because not all researchers involved in the development and application of intervention programs have access to statistical experts.

In the present paper, the EuM procedure is applied in its basic form: one treatment group is compared to one nonexperimental comparison group; the latter is considerably larger than the treatment group; the matching procedure is applied after completion of data collection. However, it can also be used under more complex conditions. For example, the EuM procedure can be applied if more than two groups need to be matched. The individuals are then not paired but arranged in groups of three or more. Furthermore, the EuM procedure can be applied before data collection is completed. This option can be useful when an intervention is observed using a larger number of observation points. In this case, we recommend matching each individual in the treatment group to more than one individual from the comparison group. Thus the damage done by drop-out in the comparison group is minimized. As an additional advantage, the EuM procedure offers the opportunity of bootstrapping if more than one matching individual is found. Bootstrapping allows one to estimate the reliability of parameter estimates (see e.g., Mooney & Duvall, 1993).

However, the proposed EuM procedure also has its limitations. First, we recommend using only a limited number of variables for matching. Larger

---

numbers can make it hard to find adequate matches. Second, to accomplish reliable matching and to overcome problems caused by drop-out, the comparison group has to be relatively large in comparison to the treatment group. This was illustrated using data from the VEL project. However, these limitations are not specific to the EuM procedure. These are limitations of matching procedures in general (see, e.g. Dehejia & Wahba, 2002).

In sum, the proposed Euclidean-distance score matching procedure seems to be a fruitful and simple method of overcoming possible bias when estimating causal treatment effects in nonexperimental designs. For future examinations of the characteristics of the EuM procedure, we recommend systematic comparisons to propensity score matching in which sample sizes, number of variables for matching, and the number of groups are varied.

---

## 5.7. References

- Berk, R.A., & Rossi, P.H. (1999). *Thinking about program evaluation* (2nd ed.). Thousand Oaks, CA: Sage.
- Birenbaum, M., & Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational sciences. *Educational Research*, 36, 93-98.
- Blalock, H. M. Jr. (1987). Some general goals in teaching statistics. *Teaching Sociology*, 15, 164-172.
- Chabris, C.F., & Glickman, M.E. (2006). Sex differences in intellectual performance. Analysis of a large cohort of competitive chess players. *Psychological Science*, 17, 1040-1046.
- Cochran, W., & Rubin, D.B. (1973). Controlling bias in observational studies. *Sankhya*, 35, 417-446.
- Cohen, P., Cohen, J., Aiken, L. S., & West, S. G. (1999). The problem of units and the circumstance for POMP. *Multivariate Behavioral Research*, 34, 315-346.
- Cox, D.R. (1992). Causality: Some statistical aspects. *Journal of the Royal Statistical Society, series A*, 155, 291-301.
- D'Agostino, Jr.R.B. (1998). Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*, 17, 2265-2281.
- D'Agostino, Jr.R.B. (2005). Propensity score. In. B.S. Everitt, & D.C. Howell (eds.), *Encyclopedia of Statistics in Behavioral Science* (pp. 1617-1619). Chichester, UK: John Wiley.
- DeGEval (2002). *Standards für Evaluation [Standards for Evaluation]*. Köln: Deutsche Gesellschaft für Evaluation e.V.
- Dehejia, R. (2005). Practical propensity score matching: a reply to Smith and Todd. *Journal of Econometrics*, 125, 355-364.
- Dehejia, R.H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84, 151-161.

- 
- Fink, A. (1995). *Evaluation for education & psychology*. Thousand Oaks, CA: Sage.
- Foster, E.M. (2003). Propensity score matching. An illustrative analysis of dose response. *Medical Care*, 41, 1183-1192.
- Gal, I., & Ginsburg, L. (1994). The role of beliefs on attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education*, 2, 1-54.
- Heckman, J.J., Ichimura, H., & Todd, P.E. (1998). Matching as an econometric evaluation estimator. *Review Economic Studies*, 65, 261-294.
- Holland, P.W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, 81, 945-960.
- Joint Committee on Standards for Educational Evaluation (JCSEE, 1994). *The program evaluation standards*, 2<sup>nd</sup> edition. Thousand Oaks, CA: Sage.
- Little, R.J.A., & Rubin, D.B. (2002). *Statistical analysis with missing data*, 2<sup>nd</sup> ed. New York: Wiley.
- Mooney, C. Z., & Duval, R. D. (1993). *Bootstrapping. A Nonparametric Approach to Statistical Inference*. Newbury Park, CA: Sage.
- Onwuegbuzie, A. J. (2000). Statistic anxiety and the role of self-perceptions. *Journal of Educational Research*, 93, 323-330.
- Pintrich, P. R., & Schunk, D.H. (1996). *Motivation in education: Theory, research and application*. Englewood Cliffs, NJ: Simon & Schuster.
- Rosenbaum, P.R., & Rubin, D.B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70, 41-55.
- Rubin, D. (1977). Assignment to a treatment group on the basis of a covariate. *Journal of Educational Statistics*, 2, 1-26.
- Schmitz, B. (2001). Self-Monitoring zur Unterstützung des Transfers einer Schulung in Selbstregulation für Studierende [Self-monitoring to support the transfer of a training to self-regulation in students]. *Zeitschrift für Pädagogische Psychologie*, 15(3/4), 181-197.
- Schafer, J.L., & Graham, J.W. (2002). Missing data: our view of the state of the art. *Psychological Methods*, 7, 147-177.
- Schober, B., Wagner, P., Reimann, R., Atria, M., & Spiel, C. (2006). Teaching research methods in an internet-based blended learning setting. *Methodology*, 2(2), 73-82.

- 
- Spiel, C. (2005). Program evaluation. In C.B. Fisher & R.M. Lerner (Eds.), *Applied developmental science: An encyclopedia of research, policies, and programs* (pp. 879-883). Thousand Oaks, CA: Sage.
- Spiel, C., Strohmeier, D., Faradji, S., Schober, B., Gradingner, P., Zens, B., Aichinger, A., & Reimann, R. (2004). Selbstreguliertes Lernen durch Vienna E-Lecturing (VEL). Konzept, Umsetzung und Evaluation [Self-regulated learning with Vienna E-Lecturing (VEL). Concept, application and evaluation]. In W. Fröhlich & W. Jütte (Eds.), *Qualitätsentwicklung in der postgradualen Weiterbildung. Internationale Entwicklungen und Perspektiven [Quality advancement in post-graduate education. International developments and perspectives]*. Münster: Waxmann.
- SYSTAT 12.0 (2007). SYSTAT Corp. <http://www.systat.com/>
- Wagner, P., Schober, B, Reimann, R., Atria, M., & Spiel, C. (in press). Vienna E-Lecturing: Trainingskonzept zum selbstregulierten Lernen im Studium [Vienna E-Lecturing: a program to enhance self-regulated learning at University]. In B. Schmitz & M. Landmann (Eds.), *Trainings zum selbstregulierten Lernen [Programs to enhance self-regulated learning]*. Stuttgart: Kohlhammer.

---

### 5.7.1. Appendix: Computational steps to create distance matrices for the EuM matching procedure

The following steps, performed using SYSTAT 12.0, create a distance matrix that can be used to establish matches. Other general purpose statistical software packages such as Splus or SPSS can easily be convinced to perform similar operations.

Standardization. Click 'Data,' 'Standardize,' select matching variables, and click 'ok.' The program replaces the original scores of the selected variables by their corresponding, standardized scores.

Sorting. Click 'Data,' 'Sort File,' and select the variables based on which the file is to be sorted. Click 'ok.' File will be sorted (re-ordered). Selecting more than one variable yields nested sorting.

Calculating distances. Click 'Analyze,' 'Correlations,' and 'Simple.' Select the variables based on which matches are created. Select 'Distance measures' and 'Euclidean.' Click 'Save matrix' and type name (plus path) for distance matrix. Click 'ok.' Distance matrix will be saved. The resulting distance matrix will, in the sample case of the present application, have the following structure.

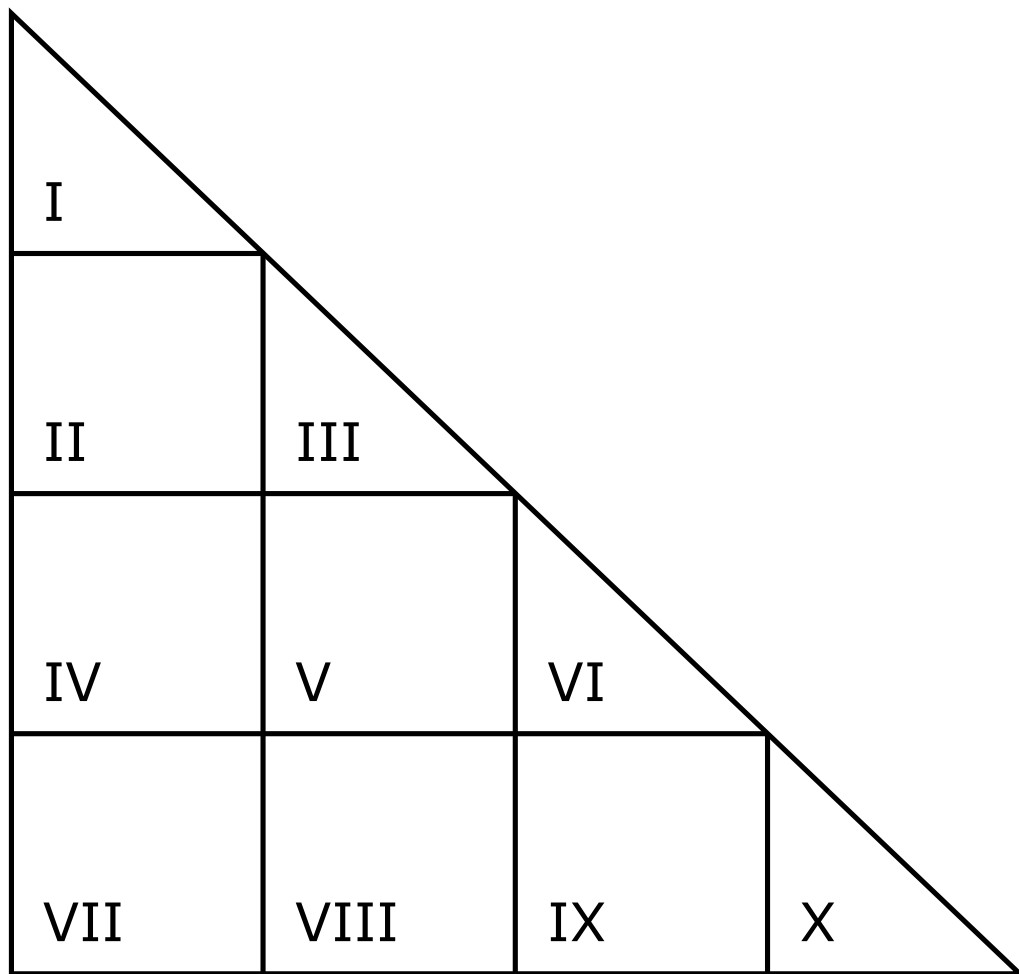


Figure A1: Structure of sorted, nested distance matrix

In the present application of the EuM procedure, the blocks of the distance matrix have the following sample interpretation:

I: Distances among women's scores in the experimental group

III: Distances among men's scores in the experimental group

VI: Distances among women's scores in the control group

X: Distances among men's scores in the control group.



---

For EuM procedure, the interesting distances are located in Blocks IV and VIII:

IV: Distances of the women in the experimental from the women in the control groups; and

VIII: Distances of the men in the experimental from the men in the control groups. In these two blocks, each column contains the distances of an individual in the experimental group to all individuals in the control group, and each row shows the distances of an individual in the control group to all individuals in the experimental group, separately for women (Block IV) and men (Block VIII).

Creating matches. In the first step, for each individual in the experimental group the best match (lowest distance) to an individual of the control group is determined. If this individual of the control group has been assigned to only one individual of the experimental group than the optimal match is identified. However, if the individual of the control group has been assigned to two individuals a second step has to be performed. To be prepared for this situation, we recommended in the first step for each individual in the experimental group not only assigning the best match but at least the best four matches. In the second step, for each of the two individuals in the experimental group, the difference between the first and the second best match (to individuals in the control group) is calculated. For the individual in the experimental group with the higher difference between the two best matches, the first match is used, whereas for the individual with the lower difference between the two best matches the second best match is used. If the lowest difference is shown for

---

more than two individuals in the experimental group this procedure has to be extended.

For matched group comparisons, each case in the experimental group for which a match was identified is assigned a 1. The matched case in the control group is assigned a 2. These scores can then be used for subsequent group comparisons in variables not used for matching.

---

## **6. Benefits of the person-oriented perspective for program-evaluation – analyzing for differential treatment effects of the Vienna E-lecturing program<sup>\*</sup>**

Dominik Lapka, Petra Wagner<sup>1</sup>, Barbara Schober, Petra Gradingner, and Christiane Spiel

Faculty of Psychology, University of Vienna, Austria

<sup>1</sup> Faculty of Applied Health and Social Sciences, Upper Austria University of Applied Sciences

<sup>\*</sup> This research was supported by the Research Fund of the Austrian National Bank, Grant No. 11107.

---

## Abstract

In program evaluation, the approaches applied to identify program effects are generally variable-oriented. This article argues for the additional utilization of person-oriented analyses to help identify differential program effects. The evaluation of a learning program (VEL-training) using a pre-post-training-control design serves to illustrate. The application of a variable-oriented perspective showed small effects with low practical relevance. Through the application of a person-oriented approach, and taking into account the theoretical framework of the intervention, three students groups which differ systematically in their learning motivation could be identified. For two of these groups the intervention was very successful. The discussion provides recommendations on how person-oriented approaches can be used systematically in program evaluation.

Key-words: person-oriented approach, program evaluation, learning motivation, variable-oriented approach

---

## 6.1. Introduction

A compliance with specific standards when performing program-evaluations is required by the scientific community in order to ensure a high level of quality. Also, most of the intentions of evaluation-customers can only be satisfied by adhering to scientific standards (cp. JCSEE, 1994), rendering them thus indispensable. One consequence of evaluations conducted under such standards is that, in general, programs lacking verifiable effects will not be implemented. One can question whether this decision is always right, as interventions are usually examined for global effects but not always for differential effects. The reasoning behind the focus on global effects is that most of the concepts applied in data analyses, and most of the hypotheses being assessed, are variable-oriented (von Eye, Bogat & Rhodes, 2006). Another reason is that, due to the common lack of prior knowledge, the formulation of differential hypotheses pertaining to the effects which can be expected from certain subgroups is rare. Consequently, existing differential effects will not be detected through global analyses based on a variable-oriented perspective. Accordingly, the intervention under evaluation will be wrongly considered to be ineffective, which in turn leads to financial losses and, of even more significance, viable educational support will not become available. This contribution, therefore, argues for the application of person-oriented analyses, in addition to the common variable-oriented approach, in order to identify the differential effects of an intervention during evaluation. In the following both approaches will be applied in the evaluation of a showcase program – an internet supported program for the promotion of self-regulated learning in an academic context (Vienna E-Lecturing – VEL). The results of the two perspectives will be compared.

---

## 6.2. Variable- vs. person-oriented approach

In order to verify the effects of a specific intervention, a summative evaluation is usually conducted which, in the simplest case, compares a treatment group to a control group. In doing so, hypotheses are generally formulated on the variable level – for example the self-efficacy of the treatment group will increase. This leads to a comparison of group means, a process which can be applied in different grades of complexity. For the most part – as in the hypotheses described – each variable will rely on one single value to describe the entire treatment group, hence the name variable-oriented approach. From a methodological perspective one assumes that, with the aggregation of data, inter-individual differences are random and thus negligible. A precondition for this approximation is the homogeneity of the treatment group with respect to the variables under analysis. Only in this case will no additional information be lost through aggregation and a better overview is thus guaranteed. However, a homogeneous treatment group is not always realistic and interventions don't always affect all individuals in the same way (Gibson, 2003; Peck, 2005). A simple and common example is the Matthew Effect, commonly found in the area of reading promotion: One instance of this effect is that good readers improve their skills while weaker pupils are over-strained and get even worse (Morgan, Farkas & Hibel, 2008). If the data compiled from a heterogeneous treatment group is aggregated, the resulting effects are biased because the general conclusions derived are not true on the individual level (von Eye, 2006). Program evaluation can thus lead to artificial results. The evaluation of a reading training can, for example, be detecting minor effects while, in reality, good pupils are

---

greatly enhancing their reading skills although the intervention is proving to have no or even negative effects on the weaker children. From a variable-oriented perspective this intervention would, at best, be rated as ineffective, although it is actually highly effective for advanced children. Alternatively, if the intervention is implemented, the weak pupils would be even more disadvantaged than before, particularly in comparison to their more advanced classmates.

To avoid such misinterpretations within program evaluation, and in order to evaluate the effectiveness of an intervention properly, the heterogeneity of the treatment group has to be taken into account. Therefore, it is essential to conduct analyses for particular individuals or homogeneous subgroups, and thus evaluate the differential effectiveness of the intervention. This is also the line of reasoning behind the person-oriented approach which has been championed in recent years by the group of researchers associated with Magnusson and Bergman as well as Spiel (Bergman & Magnusson, 1997; Magnusson, 2000; Spiel, 1998; q.v. von Eye, Bogat & Rhodes, 2006). The person-oriented approach is based upon the assumption that divergent subgroups may exist and that aggregate-level parameters may contradict parameters estimated for the entire group (von Eye, 2006). This approach was particularly promoted within the domain of developmental psychology, as individuals were starting to be seen in their integrative entirety and not just as the sum of specific variables. The basic idea was that individuals develop, not variables (Bergman, Magnusson & El-Khoury, 2003). In recent years this approach has also been adopted in other fields of psychology and the social sciences.

The current paper illustrates the advantages offered by person-oriented analyses for the purposes of evaluation. For this reason, the next section

---

presents, in short, the principles of the person-oriented approach in identifying the differential effects of an intervention. Afterwards, both variable-oriented as well as person-oriented analyses are applied, stepwise, in the evaluation of an intervention conducted in an academic context. The results of both approaches are presented and compared with each other.

### 6.3. Identifying differential effects

Often it is not known if a treatment group is heterogeneous and whether homogeneous subgroups exist within it. One possibility to check the sample for such homogeneous subgroups, and the first step in testing differential effects, is to conduct a type analysis. Here persons are merged into homogeneous groups based on scores assessing specific variables. One benefit of this approach, particularly in program evaluation, is that it is not necessary to define the different types a priori. They can be identified through exploration as well. Should this be the case, interpretations of the types discovered must be based upon a scientifically founded theory, since artificial subgroups can be found in nearly any data set (Chen, 1990; von Eye, 2006). Thus it is essential that the selection of the variables upon which the types are based, as well as the interpretation of the results, follow specific theoretical guidelines in order to avoid the identification of artificial types (Bogat, Levendosky & von Eye, 2005).

Two methods of type identification can be distinguished: clustering and categorization (von Eye & Bergman, 2003). Clustering refers to the cluster analysis methodology, which comprises several methods for finding homogeneous subgroups that are as heterogeneous as possible by focusing on specific characteristics of individuals. An exemplary application of cluster analysis within program evaluation gives Perk (2005). Using categorization Individuals



---

are merged into groups according to specific variable values. For instance sex can be coded with the first digit (1 for women and 2 for men) and age with the second (1 for young, 2 for middle aged and 3 for old); thus the number 23 stands for older men. In applying this method several variables can be categorized, and one variable can be broken down into several categories. A specific procedure using this method is the Configural Frequency Analysis (Lienert & Krauth, 1973; q.v. von Eye, Spiel & Wood, 1996), which is used to screen cross-tabulations for cells that contain significantly more cases (types) or fewer cases (anti-types) than expected from a random model. An example for categorization using this method has been published by von Eye, Bogat and Rhodes (2006).

So far few studies have systematically compared the commonly applied variable-oriented approach with the person-oriented approach (Bogat, Levendosky & von Eye, 2005; Spiel, 1998; von Eye, Bogat & Rhodes, 2006). More importantly, there is a conspicuous lack of studies to demonstrate the benefits of additional person-oriented analyses in the context of program evaluation.

In view of this scientific gap, the present paper applies and compares both approaches in the evaluation of a specific intervention – an internet supported program for the promotion of self-regulation in an academic context (Vienna E-lecturing) – in order to open new avenues towards the advancement of the quality of program evaluations. On the one hand, this paper demonstrates the benefits of additional person-oriented analyses in program evaluation, and pleads for its application. On the other hand this paper serves as manual, since the methodological procedure will be described step by step.

---

In the following section the intervention VEL (Vienna E-Lecturing) will be presented. Afterwards both methodological approaches will be specified, applied and their particular results compared with one another. Based upon this comparison, practical implications for the quality assurance of evaluations are discussed.

#### 6.4. Vienna E-Lecturing (VEL)

It is well known, and has been repeatedly demonstrated, that methods and statistics courses are not among those most popular with students majoring in Psychology. In addition, these courses are often associated with the highest levels of anxiety (Blalock, 1987; Gal & Ginsburg, 1994; Onwuegbuzie, 2000). It has been assumed that these circumstances result in negative consequences for competence development and academic progress, as indicated by poor performances in these courses. To overcome the discrepancy between relevance and acceptance of methods and statistics, the internet-based teaching concept Vienna E-Lecturing (VEL) was developed (cp. Schober et. al, 2006). It systematically combines online sessions with face-to-face lessons and imparts factual content of methods, learning competence, teamwork, and E-competence. Concerning learning competence, two areas are promoted: skill and will. While skill focuses on cognitive and meta-cognitive learning strategies as well as the competence of how to learn effectively and efficiently, will refers to the motivational and emotional components of learning competence. The promotion of the motivational aspect of learning competence is theoretically based on the model of social motivation (Dweck & Leggett, 1988), which, in addition to individual goal orientation (performance vs. competence orientation), also focuses on self-related cognitions such as the implicit personality theory, in other

---

words, the theory one holds on the flexibility or stability of one's skills and self-efficacy. Based on these variables, Dweck and Leggett (1988) explain how failure can lead either to interest or helplessness. Concretely, VEL should promote a more flexible implicit personality theory and increase competence orientation and self-efficacy so that students also increase interest, reduce helplessness and improve their factual knowledge (cp. Schober et. al, 2006).

As the evaluation of the program focuses on these motivational changes among psychology students with regard to methods and statistics, data regarding the following variables was collected: implicit personality theory, performance orientation, competence orientation, self-efficacy, interest and helplessness. Three different aspects of factual knowledge were tested: reproduction, comprehension and the production of methodological content.

The evaluation of the VEL was conducted with a treatment-control-design, whereby the VEL program participants comprised the treatment group, and the control group was formed out of the participants of a reduced program variant (in this reduced variant no supervised tasks, regular group tasks or trainings were conducted). The students decided individually, at the beginning of the academic year, which variant they wished to participate in. Thus far, data from two cohorts (n = 90; 15 males, 75 females) of treatment group participants are available, in addition to data from 375 students (50 males, 325 females) from the control group. Data collection was conducted at the beginning as well as at the end of the intervention. All motivational variables were assessed using previously published scales consisting of 3-4 items. Reliability was checked and deemed satisfactory. Details of the data collection process are available in Lapka et al. (2010).

---

## 6.5. Variable-oriented approach

The variable-oriented approach was applied to assess the general effects of the treatment. Therefore, the aggregate values (means) of the treatment group concerning the target variables are compared to those for the control group.

To confirm the effects of the VEL, it is necessary that both groups have the same aggregate value levels prior to the intervention. Otherwise, different post test values could be attributed to differences already existing prior to the intervention. To counter such systematic differences between treatment and control groups, the assignment to treatment and control group is usually randomized. Random assignment to the two versions of VEL was however impossible, because students were required to submit an application in order to participate in the full context version of the VEL. If only highly motivated students would volunteer to apply to the treatment group, estimations of the effectiveness of VEL could be biased by self-selection. Thus, in order to secure a treatment group with same starting conditions as the control group, the Euclidean-Matching (EuM) procedure (Spiel et al., 2008) was used to match control subjects to treatment students on the most relevant variables, namely *factual knowledge* in statistics, *interest* in the topic of the course, learning process *monitoring*, *organization* of the self-regulated learning process, and *time management*. The application of the EuM generated a matched control-group consisting of 84 students with no significant differences [ $F(7, 158) = 1.72, p > .05$ ] in the variables relevant to the VEL-group (also 84 students, in each case 15 males, 69 females) (Spiel et al., 2008). The means and standard deviations are depicted in Table 1.

*Table 1: Comparison of the treatment group and the control group regarding the motivational variables from the variable-oriented perspective.*

Scale	treatment group		control group	
	pre M (SD)	post M (SD)	pre M (SD)	post M (SD)
	n (84)		n (84)	
Implicit personality theory	4.82 (0.75)	5.15 (0.69)	4.89 (0.76)	5.08 (0.74)
Competence orientation	4.59 (0.79)	4.56 (0.78)	4.43 (0.82)	4.33 (0.83)
Performance orientation	3.42 (1.03)	3.56 (1.12)	3.53 (0.93)	3.64 (1.06)
Self-efficacy	4.46 (0.71)	4.80 (0.63)	4.68 (0.71)	4.84 (0.64)
Interest	3.93 (1.07)	4.15 (1.03)	3.90 (0.89)	3.91 (1.00)
Helplessness	2.90 (1.00)	2.35 (0.87)	2.59 (1.09)	2.25 (0.92)

First the development of the two groups across the motivational variables is described by the interactions found with two-way ANOVAs (treatment vs. control group; pre- vs. post-test), which are presented in Table 2. These interactions provide information concerning the progress of the treatment group relative to the control group (intergroup). Small effects regarding self-efficacy and interest were identified, whereas the treatment group showed better development than the control group (see Table 2).

*Table 2: Variable-oriented results of the intergroup effects between the treatment and control group at two points in time*

Scale	<i>F</i>	<i>df</i>	<i>p</i>	<i>D</i>
Implicit personality theory	1.38	1, 165	>.05	.19
Competence orientation	0.51	1, 165	>.05	.09
Performance orientation	0.04	1, 167	>.05	.03
Self-efficacy	4.19	1, 167	<.05	.27
Interest	3.66	1, 167	<.05	.21
Helplessness	2.13	1, 165	<.10	-.22

*Note:* The table contains the results of the interaction of the two-way ANOVA (treatment vs. control group x pre vs. post-test) (*F*, *df*, *p* and the effect-size *D*).

Information concerning the absolute development of the treatment group in comparison to the control group is provided by the second step, the intra-group analyses. To this end the differences between the pre-test and post-test values

were analysed separately for each motivational variable with t-tests for both the treatment and control group (see Table 3).

From a variable-oriented perspective, one sees a positive development in both the treatment and the control group regarding implicit personality theory (increment), self-efficacy and helplessness. The effect sizes for the treatment group are moderate whereas the effect sizes found in the control group are small (see Table 3). Interest only increases in the treatment group with a small effect size.

*Table 3: Variable oriented perspective of the intra-group-effects concerning pre and post test measurements*

Scale	treatment group				control group			
	<i>t</i>	<i>df</i>	<i>p</i>	<i>D</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>D</i>
Implicit personality theory	3.97	82	<.001	.46	2.47	82	<.01	.25
Competence orientation	0.36	82	>.05	-.04	-1.25	82	>.05	-.12
Performance orientation	1.43	83	<.10	.13	1.32	83	<.10	.11
Self-efficacy	5.54	83	<.001	.51	2.75	83	<.01	.24
Interest	2.63	83	<.01	.21	0.08	83	>.05	.01
Helplessness	-5.96	82	<.001	-.59	-2.88	82	<.01	-.34

*Note:* The table contains the two-tailed t-tests (*t*, *df*, *p*) as well as the effect size *D*. The scales range from 1 (= completely disagree) to 6 (= completely agree).

Thirdly, the intergroup effects of the intervention regarding factual knowledge were assessed by conducting a MANOVA, using previous knowledge as the covariate. Factual knowledge at the end of the intervention was tested with regard to three different aspects: reproduction, comprehension, and production.

Results show a multivariate effect [ $F(3, 162) = 3.885, p < .05$ ] that is expressed on univariate level through differences in productive knowledge [ $F(1, 164) = 7.619, p < .01, d=0.34$ ], whereas the treatment group shows higher values ( $M_{treatment} = 64.21; M_{control} = 56.58$ ).

---

To sum up, the variable-oriented approach discloses little benefits for the treatment group in contrast to the control group. Both groups develop a more flexible implicit personality theory, gain in self-efficacy and reduce their helplessness. The treatment group increased more in self-efficacy and interest than the control group. Although these interactions are significant, their practical relevance (the effect sizes) are rather small. Also the effect size of the significant difference regarding productive knowledge between the two groups is low.

#### 6.6. Person-oriented approach

By application of the person-oriented approach one assumes that, in the sample presented, different motivational subgroups will prove to exist. In this case, the aggregate values of the variable-oriented approach will not truly represent the different patterns and changes of the individuals and/or specific subgroups. However, these differences may influence the effect of the intervention, for example the intervention may have a higher effect on motivated students and no effect on disinterested students or vice versa.

In order to obtain information concerning the differential effectiveness of the VEL program, the person-oriented approach was applied in three steps. First, different motivational subgroups were identified through the application of a cluster analysis and interpreted with respect to the theoretical model of the intervention. Second, the distribution of these motivational subgroups in the treatment and control groups was checked for uniformity, to ensure that treatment and control group share the same baseline. The third step proceeds with differential analyses of learning competence (*intergroup* and *intra-group*

---

analyses) and factual knowledge (*intergroup* analyses). Analyses must be done separately for each motivational type.

In accordance with the theoretical background of the VEL, the central variables of the social motivation model (Dweck & Leggett, 1988; see before) were considered in the cluster analysis: flexible implicit personality theory, competence orientation, performance orientation and self-efficacy. The cluster analysis was applied in three steps: the single linkage method to identify outliers, the hierarchical ward-method and the non-hierarchical k-means to optimize cluster assignment (Lapka et al., 2010). Three motivational types could be identified at the beginning of the intervention: *motivationally balanced* students, *competence oriented* students and *students with motivational deficits*. The *motivationally balanced* students show high values in all four variables. The *competence oriented* students show the same characteristics as the *motivationally balanced* with the exception of lower values regarding performance orientation. The *students with motivational deficits* have a less flexible implicit personality theory, are less competence oriented, show poorer values regarding self-efficacy and are thus more helpless and less interested (for more detail see Lapka et al. in press). In order to avoid artificial subgroups, a close reliance on a theoretical framework is necessary for both variable selection as well as the interpretation of the different motivational types.

As was previously mentioned, the students (three motivational subgroups) could voluntarily decide whether or not to attend the VEL program or just participate in the reduced version of VEL (and serve as controls). Therefore in the second step, using a  $\chi^2$ -Test, we tested for the existence of systematic differences on the basis of self-selection. Results show that the three



motivational types are distributed uniformly in the treatment and control groups [ $\chi^2(2, N = 430) = 3.31, p = .19$ ] (see Table 4).

Although Wilk's Lambda ( $\Lambda=0.160$ ) suggests that each motivational type is relatively homogeneous, each cluster was checked via MANOVA for initial differences between treatment and control group. Differences were found to exist solely among the *students with motivational deficits* [ $F(7, 138) = 2.52, p < .05$ ]. This is manifested on the univariate level by differences in interest [ $F(1, 144) = 6.18, p < .05, d=0.47$ ], whereas students with higher interest are rather assigned to the treatment group (see Table 4).

*Table 4: Comparison of the treatment group and the control group regarding the motivational variables from the person-oriented perspective.*

Scale	treatment group		control group	
	pre M (SD)	post M (SD)	pre M (SD)	post M (SD)
motivational balanced students				
	n (26)		n (117)	
Implicit personality theory	5.09 (0.52)	5.15 (0.66)	5.17 (0.50)	5.17 (0.67)
Competence orientation	4.79 (0.63)	4.62 (0.91)	4.57 (0.73)	4.31 (0.77)
Performance orientation	4.27 (0.74)	4.23 (1.07)	4.35 (0.65)	4.29 (0.83)
Self-efficacy	4.80 (0.44)	4.99 (0.46)	4.96 (0.45)	4.99 (0.57)
Interest	4.12 (0.80)	4.16 (0.98)	3.86 (0.86)	3.78 (1.05)
Helplessness	2.65 (0.88)	2.31 (0.83)	2.35 (0.87)	2.21 (0.87)
competence oriented students				
	n (33)		n (91)	
Implicit personality theory	5.18 (0.59)	5.45 (0.54)	5.28 (0.57)	5.31 (0.63)
Competence orientation	4.83 (0.68)	4.81 (0.67)	4.79 (0.58)	4.53 (0.73)
Performance orientation	2.39 (0.55)	2.81 (0.93)	2.45 (0.56)	3.04 (0.89)
Self-efficacy	4.77 (0.47)	4.97 (0.62)	4.99 (0.57)	5.04 (0.59)
Interest	4.20 (1.07)	4.44 (0.95)	4.12 (0.84)	4.03 (1.01)
Helplessness	2.58 (1.00)	2.01 (0.84)	2.27 (0.89)	2.07 (0.91)
students with motivational deficits				
	n (26)		n (109)	
Implicit personality theory	4.06 (0.60)	4.76 (0.74)	3.85 (0.61)	4.35 (0.80)
Competence orientation	4.10 (0.84)	4.20 (0.66)	4.02 (0.66)	3.98 (0.68)
Performance orientation	3.86 (0.71)	3.75 (0.93)	4.03 (0.76)	3.96 (0.93)
Self-efficacy	3.71 (0.68)	4.39 (0.62)	3.91 (0.63)	4.22 (0.58)
Interest	3.38 (1.10)	3.78 (1.09)	2.95 (0.70)	3.23 (0.85)
Helplessness	3.63 (0.79)	2.81 (0.79)	3.86 (0.93)	3.20 (0.95)

*Note:* The table contains means (M) and standard deviations (SD) of the two groups from the global perspective as well as from the type-specific perspective, using the three motivational types. The data regard the pre-test at the beginning of VEL and at the post-test at the end of the first term. The scales range from 1 (= completely disagree) to 6 (= completely agree).

In the third step, the differential effects of the intervention were assessed. To this end the progress concerning motivational aspects (*intergroup* and *intra-group* analyses) as well as factual knowledge (*intergroup* analysis) was analysed separately for each motivational type.

First the *interactions* between treatment and control-group were computed for each type with a two-way ANOVA (treatment vs. control group; pre- vs. post-test). The results of these *interactions* are presented in Table 5. The *motivationally balanced students* show no effect in any motivational variable. Treatment and control groups show the same progress patterns. Medium effects in interest and helplessness were found among the *competence oriented students*, whereas the treatment group shows more positive development than the control group. Also, medium effects favoring the treatment group were found among the *students with motivational deficits* in self-efficacy (see Table 5).

Again, these *interactions* only provide information concerning the progress of the treatment group relative to the control group (*intergroup*). In order to obtain information concerning the absolute development of the treatment in contrast to the control group, *intra-group* analyses were performed for each motivational type. Their results confirm differences in progress among the three motivational types (see Table 6).

*Table 5: Person-oriented results of the intergroup effects between the treatment and control group at two points in time for each motivational type.*

Scale	<i>F</i>	<i>df</i>	<i>p</i>	<i>D</i>
motivationally balanced students				
Implicit personality theory	0.040	1, 130	.842	0.107
Competence orientation	0.140	1, 130	.709	0.104
Performance orientation	0.089	1, 130	.766	0.027
Self-efficacy	2.307	1, 130	.131	0.344
Interest	0.401	1, 130	.528	0.144
Helplessness	0.963	1, 130	.328	-0.241

competence oriented students				
Implicit personality theory	3.548	1, 107	.062	0.417
Competence orientation	3.491	1, 107	.064	0.356
Performance orientation	0.998	1, 107	.320	-0.225
Self-efficacy	1.758	1, 107	.188	0.280
Interest	6.332	1, 107	.013	0.345
Helplessness	5.637	1, 107	.019	-0.402
students with motivational deficits				
Implicit personality theory	1.826	1, 122	.179	0.294
Competence orientation	0.875	1, 122	.351	0.190
Performance orientation	0.006	1, 122	.938	-0.034
Self-efficacy	9.230	1, 122	.003	0.588
Interest	0.987	1, 122	.322	0.134
Helplessness	0.631	1, 121	.429	-0.175

Note: The table contains the results of the interaction of the two-way ANOVA (treatment vs. control group x pre vs. post-test) ( $F$ ,  $df$ ,  $p$  and the effect-size  $D$ ).

During the intervention, the *motivationally balanced* students in the treatment group show moderately positive effects in gaining self-efficacy and reducing helplessness. In contrast to these positive effects, the control group became less competence orientated after the intervention. More positive development can be found among the *competence oriented* students. The students in the treatment group show positive changes with medium effect sizes in all variables. Concretely, they acquired a more flexible implicit personality theory, increased their performance orientation and reduced their helplessness. In contrast, the students in the control group developed less positively. They reduced their competence orientation and increased their performance orientation. The controls also reduced their helplessness, but the effect size is much lower. The most positive changes linked with the highest effect sizes are found among the *students with motivational deficits*. Members of both the control and treatment groups obtained a more flexible implicit personality theory, gained in self-efficacy and interest, and reduced their helplessness. The effect sizes in

the treatment group are, with the exception of interest, much higher than those for the control group.

*Table 6: Type-specific perspective of the intra-group-effects*

Scale	treatment group				control group			
	<i>t</i>	<i>df</i>	<i>p</i>	<i>D</i>	<i>t</i>	<i>df</i>	<i>p</i>	<i>D</i>
motivationally balanced students								
Implicit personality theory	-0.451	25	.656	0.108	-0.430	105	.668	0.001
Competence orientation	1.142	25	.264	-0.229	3.301	105	.001	-0.346
Performance orientation	0.207	25	.838	-0.038	1.142	105	.256	-0.078
Self-efficacy	-3.004	25	.006	0.429	-0.397	105	.692	0.051
Interest	-0.304	25	.764	-0.054	0.807	105	.421	-0.090
Helplessness	2.169	25	.040	-0.405	1.482	105	.141	-0.158
competence oriented students								
Implicit personality theory	-2.398	32	.022	0.483	-0.565	75	.574	0.051
Competence orientation	0.178	32	.860	-0.030	3.496	75	.001	-0.391
Performance orientation	-2.782	32	.009	0.554	-6.051	75	<.001	0.796
Self-efficacy	-1.896	32	.067	0.364	-0.836	75	.406	0.071
Interest	-1.914	32	.065	0.239	1.803	75	.075	-0.100
Helplessness	3.871	32	.001	-0.612	2.127	75	.037	-0.223
students with motivational deficits								
Implicit personality theory	-4.301	24	<.001	1.039	-6.416	98	<.001	0.688
Competence orientation	-0.857	24	.400	0.129	0.499	98	.619	-0.056
Performance orientation	0.387	24	.702	-0.128	1.031	98	.305	-0.092
Self-efficacy	-5.481	24	<.001	1.041	-5.186	98	<.001	0.510
Interest	-2.386	24	.025	0.370	-3.014	98	.003	0.357
Helplessness	4.588	24	<.001	-1.032	5.609	97	<.001	-0.703

*Note:* The table contains the two-tailed t-tests (*t*, *df*, *p*) as well as the effect size *D*. The scales range from 1 (= completely disagree) to 6 (= completely agree).

Finally, the *intergroup* effects of the intervention regarding factual knowledge were verified with a MANOVA, using previous knowledge as a covariate, for each of the three motivational types. There were no differences in factual knowledge between the treatment and the control groups among the *motivationally balanced* students [ $F(3, 138) = 1.524, p > .05$ ] and the *competence oriented* students [ $F(3, 119) = 0.549, p > .05$ ]. However, among the *students with motivational deficits*, there were multivariate differences between the treatment and control groups after the intervention [ $F(3, 138) = 3.776, p < .05$ ], and they are expressed on univariate level through differences in productive knowledge

---

[ $F(1, 132) = 9.747, p < .01, d=0.786$ ]. The treatment group shows higher percentage values ( $M_{treatment} = 69.87; M_{control} = 54.10$ ).

In summation, the person-oriented approach reveals that there is one group of students which does not benefit from the VEL program at all: the *motivationally balanced students*. These are highly motivated learners who, in addition to a high competence orientation, also display high performance orientation. For these students there was nearly no difference between those who participate in the VEL or the control group. The most relevant differences in progress between treatment and control group were found among the *competence oriented students*, who are also highly motivated but not very performance oriented. Students of this type found in the treatment group raised their interest levels and reduced their feelings of helplessness concerning research methods more than the corresponding controls. *Students with motivational deficits* benefit most. They had the highest backlog concerning motivational variables at the beginning of the intervention. Both the treatment and control group members show high positive changes regarding implicit personality theory, self-efficacy, interest and helplessness. Nevertheless, the students in the treatment group show an even better development during VEL program, particularly with respect to self-efficacy. Additionally, the treatment group showed a higher factual knowledge of research methods concerning productive content after the intervention. Thus, the person-oriented approach reveals a differentiated impact of the intervention. While one group of students benefited immensely, another group amassed few if any advantages.

In order to make the program more efficient one must consider omitting *motivationally balanced* students from the VEL, since the control condition

---

provided these students with equivalent benefits. Additionally, the intervention should either (1) focus on only the *competence oriented* students, which leads to high overall benefits, as *students with motivational deficits* also benefit significantly from the reduced version, or (2) the VEL should be used in terms of a secondary prevention for *students with motivational deficits*, which will accordingly boost the motivational aspects of their learning competence. Also a combination of these two alternatives can be considered.

## 6.7. Summary and Outlook

The goal of this paper was to demonstrate the benefits of person-oriented analyses in program evaluation. For the purposes of demonstration both approaches – the variable-oriented and person-oriented approaches – were applied in the evaluation of the Vienna E-Lecturing program. From the variable-oriented perspective, the effects of the intervention could be confirmed even though the practical relevance was rather low. Based on the theoretical frame of the intervention, and using the person-oriented perspective, three different motivational types were identified within the heterogeneous sample. With respect to these three types, a differential impact of the program could be shown. Concretely speaking, the VEL shows a major practical relevance regarding training goals among competence oriented students as well as students with motivational deficits – the effect sizes for these two types, as measured by changes after participating in the VEL, are moderate to high. Motivationally balanced students, on the other hand, didn't benefit from the VEL at all.

These results illustrate how changing the perspective in data analyses – namely from the common variable-oriented perspective to a focus on the individual – leads to more precise conclusions concerning the evaluated

---

intervention, in particular it reveals possible differential effects. This translates into an enhancement of the quality of the evaluation as well as its utility for stakeholders (Patton, 1996). Based on these findings, the target group of the intervention can be confined, which in turns results in a higher efficiency of the intervention.

In order to facilitate the application of person-oriented analyses in program evaluation, criteria have to be defined. As shown by this study, a good theoretical foundation is essential for type identification, otherwise artificial results are a distinct possibility (Bogat, Levendosky & von Eye, 2005; von Eye, 2006). Also, it is necessary for the different subgroups to be represented equally in treatment and control groups. Otherwise the assignment to control vs. treatment group is confounded by types, and the effects found cannot be ascribed solely to the intervention. Additional cross validation can be taken into consideration in order to ensure true effects. When defining such criteria, positive expert knowledge regarding person-oriented analyses in the fields of developmental psychology and psychopathology can be used (Bergman, Magnusson & El-Khoury, 2003).

Person-oriented analyses in program evaluation are particularly beneficial in evaluating interventions that take place in real settings. In these contexts, it is rare to encounter optimal conditions when implementing an evaluation (Fagan & Mihalic, 2003; Greene, Benjamin & Goodyear, 2001; Lucke, Donald, Dower & Raphael, 2001). The effects of the intervention are thus more difficult to determine than they would be under controlled laboratory conditions. Therefore, the limited possibilities available to evaluators, in the form of sub-optimal basic conditions, have to be compensated for by insuring methodological strengths. An

---

example of such compensation is the mixed-methods approach (Chen, 2006; q.v. Spiel, Strohmeier & Atria, 2008) – thus the systematic combination of qualitative and quantitative methods. Another example is the euclidean distance-based matching procedure (Spiel et al., 2008), which adjusts between-group comparisons for pre-existing differences if randomization is difficult or impossible to achieve. Also, this paper offers potential for methodological compensation in real life settings, particularly when the evaluator cannot influence the general set-up. As shown in this paper, the person-oriented perspective is an efficient and reliable way to obtain results, which are both more precise and more useful.



---

## 6.8. References

- Bergman, L.R. & Magnusson, D., 1997. A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology*, 9, pp. 291-319.
- Bergman, L.R. Magnusson, D. & El-Khoury, B.M., 2003. *Studying individual development in an interindividual context. A person-oriented approach*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Blalock, H.M.Jr., 1987. Some general goals in teaching statistics. *Teaching Sociology*, 15, pp. 164-172.
- Bogat, G.A. Levendosky, A.A. & von Eye, A., 2005. The future of research on intimate partner violence – person-oriented and variable-oriented perspectives. *American Journal of Community Psychology*. 36, pp. 49-70.
- Chen, H., 1990. *Theory-driven evaluations*. Newbury Park, CA: Sage.
- Chen, H., 2006. A theory-driven evaluation perspective on mixed methods research. *Research in the schools*, 13, pp. 75-83.
- Dweck, C.S. & Leggett, E.L., 1988. A social-cognitive approach to motivation and personality. *Psychological Review*, 95, pp. 256-273.
- Fagan, A. & Mihalic, S., 2003. Strategies for enhancing the adoption of school-based prevention programs – lessons learned from the blueprints for violence prevention replications of the life skills training program. *Journal of Community Psychology*, 31, pp. 235-253.
- Gal, I. & Ginsburg, L., 1994. The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education*, 2, pp. 1-54
- Greene, J. Benjamin, L. & Goodyear, L., 2001. The merits of mixing methods in Evaluation. *Evaluation*, 7, pp. 25-44.
- Joint Committee on Standards for Educational Evaluation. 1994. *The program evaluation standards* (2nd ed.). Thousand Oaks, CA: Sage.
- Lapka, D. Wagner, P. Schober, B. Grading, P. & Spiel, C. (2010). Methodenlehre: Alptraum oder Herausforderung für Psychologiestudierende? Eine Typologie auf Basis des sozialkognitiven Motivationsmodells von Dweck, [Methodology classes:

---

Nightmare or a challenge for Psychology students? A typology based on Dweck's social-cognitive approach to motivation]. *Psychologie in Erziehung und Unterricht*.

- Lienert, G.A. & Krauth, J., 1973. *KFA. Die Konfigurationsfrequenzanalyse und ihre Anwendung in Psychologie und Medizin*, [The Configural Frequency Analysis and its application in Psychology and Medicine]. Freiburg: Alber.
- Lucke, J. Donald, M. Dower, J. & Raphael, B., 2001. Considerations in the design of a mixed-method cluster. *Evaluation*, 7, pp. 110-131.
- Magnusson, D., 2000. The individual as the organizing principle in psychological inquiry: A holistic approach. In: L.R. Bergman, R.B. Cairns, L.G. Nilsson & L. Nystedt, eds 2000. *Developmental science and the holistic approach*. Mahwah, NJ: Erlbaum, pp.33-47.
- Morgan, P.L. Farkas, G. & Hibel, J., 2008. Matthew effects for whom? *Learning Disability Quarterly*, 31, pp. 187-198
- Onwuegbuzie, A.J., 2000. Statistics anxiety and the role of self-perceptions. *The Journal of Educational Research*, 93, pp. 323-330.
- Patton, M.Q., 1996. *Utilization focused Evaluation* (3th ed.). Thousand Oaks, CA: Sage.
- Schober, B. Wagner, P. Reimann, R. Atria, M. & Spiel, C., 2006. Teaching Research Methods in an Internet-Based Blended-Learning Setting. *Methodology*, 2, 73-82.
- Spiel, C., 1998. Four methodological approaches for the study of stability and change in development. *Methods of Psychological Research - online* 3 (2), <http://www.ppm.ipn.uni-kiel.de/mpr/issue5/art2/article.html> (acessed November 12, 2001).
- Spiel, C. Strohmeier, D. & Atria, M., 2008. Von der notwendigen Verschränkung quantitativer und qualitativer Methoden in der Programmevaluation am Beispiel des WiSK - Soziales Kompetenztraining für Schülerinnen und Schüler. In: F. Hofmann, C. Schreiner and J. Thonhauser eds. 2008. *Zur Komplementarität von qualitativen und quantitativen Aspekten in der erziehungswissenschaftlichen Forschung*, [On the need to offset quantitative and qualitative methods in program evaluation using WiSK as an example - Social competence training for male and female pupils]. Münster: Waxmann, pp.263-280.

- 
- Spiel, C. Lapka, D. Grading, P. Zodlhofer, E.M. Reimann, R. Schober, B. Wagner, P. & von Eye, A., 2008. A euclidean distance-based matching procedure for non-randomized comparison studies. *European Psychologist*, 13 (3), pp. 180-187.
- von Eye, A., 2006. Person-Oriented and Variable-oriented Research: Concepts, results, and development. *Merrill-Palmer Quarterly*, 52, pp. 390-420.
- von Eye, A. & Bergman, L.R., 2003. Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology*, 15, pp. 553-580.
- von Eye, A. Bogat, G.A. & Rhodes, J.E., 2006. Variable-oriented and person-oriented perspectives of analysis: The example of alcohol consumption in adolescence. *Journal of Adolescence*, 29, pp. 981-1004.
- von Eye, A. Spiel, C. & Wood, P.K., 1996. CFA models, tests, interpretation, and alternatives - a rejoinder. *Applied Psychology*, 45, pp. 345-352.

---

## **7. Methodenlehre: Alptraum oder Herausforderung für Psychologiestudierende? Eine Typologie auf Basis des sozialkognitiven Motivationsmodells von Dweck\***

Methodology: Nightmare or Challenge for Psychology Undergraduate Students? A Typology based on Dweck's Social-Cognitive Motivation Model

Lapka D., Wagner P.<sup>1</sup>, Schober B., Grading P., Reimann R. & Spiel C. (2010). Methodenlehre: Alptraum oder Herausforderung für Psychologiestudierende? Eine Typologie auf Basis des sozialkognitiven Motivationamodells von Dweck. *Psychologie in Erziehung und Unterricht*, 57 (3).

Fakultät für Psychologie, Universität Wien, Österreich

<sup>1</sup> Fakultät für Gesundheit und Soziales, Fachhochschule Oberösterreich, Österreich

\* Dieses Projekt wurde gefördert vom Jubiläumsfonds der Österreichischen Nationalbank (Projektnummer 11107).

---

## Summary

The study describes (1) systematically the motivational situation of psychology students concerning the subject research-methods using the social-cognitive motivation model (Dweck & Leggett, 1988) and proves (2) the central assumptions of the model. In a cross section design data were collected from 465 psychology students of the University of Vienna. On average, results show low subject knowledge and suboptimal values in interest and helplessness of the students. All observed gender differences (in knowledge, helplessness and implicit personality theory) are in favour of male students. The cluster analysis showed three motivational types, which in comparison support the central assumptions of the model of Dweck and Leggett (1988): the "stable negative oriented students" (type 3) show poorer values in subject knowledge and all other variables than the "competence -" (type 1) and the "achievement oriented students" (type 2). Recommendations for interventions and for an extension of the model were made.

key words: learning motivation, implicit theories, academic achievement, methodology, cluster analysis

---

## Zusammenfassung

Die vorliegende Studie beschreibt (1) unter Verwendung des sozialkognitiven Motivationsmodells (Dweck & Leggett, 1988) die motivationale Situation Psychologiestudierender im Fach Methodenlehre und prüft (2) die zentralen Annahmen dieses Modells. In einer Querschnittsuntersuchung wurden dazu 465 Studierende (400 Frauen und 65 Männer) des Diplomstudiums Psychologie an der Universität Wien schriftlich befragt und getestet. Die Ergebnisse zeigen ein insgesamt geringes Methodenwissen der Studierenden sowie suboptimale Ausprägungen in den Variablen Interesse und Hilfslosigkeit. Männliche Studierende weisen höheres Wissen, geringere Hilfslosigkeit und günstigere Ausprägungen der impliziten Persönlichkeitstheorie auf als weibliche Studierende. Eine Clusteranalyse ergab drei Motivations-Typen, deren Vergleich die zentralen Annahmen des Modells von Dweck und Leggett (1988) bestätigt: die "Stabil negativ orientierten Studierenden" (Typ 3) schneiden im Fachwissen sowie in allen weiteren untersuchten Variablen schlechter ab als die "Kompetenzorientierten Studierenden" (Typ 1) und die "Leistungsorientierten Studierenden" (Typ 2). Schlussfolgernd werden Förderansätze abgeleitet sowie eine Erweiterung des Modells empfohlen.

Schlüsselbegriffe: Lernmotivation, Implizite Theorien, Leistung, Methodenlehre, Clusteranalyse

---

## 7.1. Einleitung

Im Psychologiestudium zeigt sich oft eine Diskrepanz zwischen der Relevanz der Methodenausbildung und ihrer Akzeptanz seitens der Studierenden. Viele Studierende sozialwissenschaftlicher Studienfächer erleben die Methodenausbildung als "terminal activity". Darunter versteht man, dass die Inhalte der Methodenlehre den Studierenden zwar irrelevant erscheinen und sie nach der Ausbildung nichts mehr mit Methodenlehre zu tun haben wollen, sie jedoch durch den Studienplan zu einer kurzfristigen Auseinandersetzung mit dem Thema gezwungen sind (Blalock, 1987).

Zur Problemlage der Methodenausbildung im Psychologiestudium liegen bereits einige Studien vor. Dabei reicht das Spektrum der untersuchten Variablen von der allgemeinen Einstellung, dem Interesse und der Angst vor Statistik bis hin zur erlernten Hilflosigkeit. Es gibt jedoch nur wenige Studien, die diese Variablen zusammenhängend und auf Basis eines theoretischen Modells analysieren. Dies scheint jedoch für eine Klärung der hier wirksamen Mechanismen sinnvoll und notwendig.

Vor diesem Hintergrund verfolgt die im Folgenden dargestellte Studie zwei Forschungsziele. (1) Zunächst ist es Anliegen der vorliegenden Studie, die motivationale Situation von Studierenden der Methodenlehre (als Teil ihres Psychologiestudiums) im Kontext des sozialkognitiven Motivationsmodells von Dweck und Leggett (1988) zu untersuchen. Die bisherige Forschung belegt eindeutig die Bedeutsamkeit der im Modell beschriebenen Motivationskomponenten für das Verhalten in Leistungssituationen. Allerdings standen meist nur einzelne Komponenten des Modells im Vordergrund des Interesses (vgl. z.B. Heller, Finsterwald & Ziegler, 2001). (2) Darauf aufbauend

---

ist es das zweite Anliegen der Studie, auf einer Metaebene zu prüfen, ob die erhobenen Daten das sozialkognitive Motivationsmodell stützen. Dies wäre ein erster notwendiger Schritt zur komplexeren Prüfung des Modells in seiner Gesamtheit.

## 7.2. Psychologiestudierende und Methodenausbildung

Der Methodenausbildung kommt im Psychologiestudium ein hoher Stellenwert zu. Die Relevanz fundierter Methodenkenntnisse ist nicht nur für das Durchführen eigener, sondern grundsätzlich für das Verstehen und Interpretieren empirischer Studien erforderlich. Die Bedeutung der Methodenausbildung innerhalb des Psychologiestudiums lässt sich auch sehr gut an den Studienplänen erkennen. So entfallen z.B. im Diplomstudienplan Psychologie an der Universität Wien von den insgesamt 135 Pflichtsemesterwochenstunden in Psychologie 22 (16%) auf die Fächer "Methodenlehre" sowie "Forschungsmethoden und Evaluation". Zählt man das Fach "Psychologische Diagnostik" mit 14 Wochenstunden ebenfalls dazu, entfallen insgesamt 27 Prozent, d.h. mehr als ein Viertel des gesamten Psychologiestudiums auf die Methodenausbildung. Auch an vielen deutschen Universitäten nimmt die Methodenlehre v.a. im Grundstudium einen umfangreichen Platz ein und kann in den ersten beiden Semestern sogar bis zur Hälfte der Studienzeit in Anspruch nehmen. Dennoch zweifeln viele Studierende an der Notwendigkeit und Nützlichkeit von Statistik und Methodenlehre (Diehl, 1993). Die von Studierenden genannten Zweifel unterteilen Stark und Mandl (2000) in vier inhaltliche Kategorien: (1) Probleme der Generalisierbarkeit empirischer Ergebnisse ("*Surveys never concern the current situation*"), (2) die Reduktion von Komplexität ("*Results show only an*



---

*average view of things*"), (3) die Subjektivität und Manipulierbarkeit empirischer Daten (*"There is no sufficient data protection"*) und (4) eine zu geringe praktische Relevanz (*"Too little reference to issues with practical relevance"*). Angesichts dieser Skepsis gegenüber Statistik und Methodenlehre verwundert es nicht, dass Studierende im Laufe des Studiums zunehmend das Interesse an diesem Teil der Ausbildung verlieren (Giesler, 1998).

Häufig zeigen sich darüber hinaus sogar deutlich aversive Reaktionen, wie Gefühle der Ablehnung und des Unwohlseins im Umgang mit methodischen Inhalten (Diehl, 1993). Immerhin 70 bis 80 Prozent der Studierenden leiden unter starkem Unbehagen und Statistik-Angst (Birenbaum & Eylath, 1994; Onwuegbuzie, 2004; Onwuegbuzie & Wilson, 2003; Zeidner, 1991), weshalb Gruber und Renkl (1996) von einem "Alptraum für Studierende" sprechen.

Die Angst vor Statistik wird offenbar von weiblichen Studierenden gravierender erlebt als von männlichen, denn mehrheitlich zeigen die bisher durchgeführten Studien Geschlechtsunterschiede zu Ungunsten der weiblichen Studierenden (z.B. Benson, 1989; Roberts & Saxe, 1982), wenngleich Geschlechtsunterschiede nicht durchgängig beobachtet werden (z.B. Cherian & Glencross, 1997; siehe zusammenfassend Hilton, Schau & Olsen, 2004). Dies kann zum Einen – ähnlich wie in der Koedukationsdebatte – im Sinne der hohen Relevanz der konkreten Unterrichts- und Instruktionsgestaltung interpretiert werden (vgl. z.B. Dresel, Stöger & Ziegler, 2005). Zum Anderen stellt sich allerdings bei einigen Studien aufgrund methodischer Mängel auch die Frage nach der Validität der Ergebnisse. Die Studie von Cherian und Glencross (1997) beispielsweise beruht auf einer Stichprobengröße von nur 38 Studierenden (18 männlich, 20 weiblich). Auch Onwuegbuzie und Wilson (2003) kritisieren die

---

Validität der verwendeten Messinstrumente in diesem Bereich. D.h. insgesamt ist damit eine eher heterogene Befundlage hinsichtlich der Frage von Geschlechtsunterschieden verbunden mit einer generellen methodischen Schwäche einiger Studien in diesem Bereich erkennbar.

Jenseits der Frage von Geschlechtsunterschieden wird die Statistik-Angst häufig als "Gegenpol" zur mathematischen Selbstwirksamkeit gesehen (Onwuegbuzie, 2000), womit sich auch der Bogen zur motivationalen Situation der Studierenden spannt. Noch deutlicher wird die Relevanz der motivationalen Determinanten vor dem Hintergrund jener Befunde, die zeigen, dass sich eine positive Motivationslage günstig auf die Leistungsfähigkeit in Statistik auswirkt (Onwuegbuzie & Wilson, 2003; Roberts & Saxe, 1982). Positive Zusammenhänge mit der Leistung zeigen sich bei zahlreichen motivationalen Determinanten, wie beispielsweise der Selbstwirksamkeit (Multon, Brown & Lent, 1991; Onwuegbuzie, 2003), der flexiblen impliziten Persönlichkeitstheorie (Dweck & Molden, 2005), dem Interesse (Schiefele, Krapp & Schreyer, 1993) oder dem Flow-Erleben (Engeser, Rheinberg, Vollmeyer & Bichoff, 2005). Durch diese Befunde wird die Relevanz von gezielter Motivationsförderung im Bereich Methodenlehre evident, umso mehr da diese von Statistik-Lehrenden oft vernachlässigt wird (Gal & Ginsburg, 1994).

### 7.3. Alptraum oder Herausforderung – motivationspsychologische Determinanten

Motivation ist ein komplexes Konstrukt, das auch keineswegs einheitlich definiert wird (vgl. zusammenfassend Schober, 2002). Ein aktueller Ansatz der Motivationsforschung (siehe auch Pintrich & Schunk, 1996) geht jedoch davon aus, dass unter Motivation "...diejenigen psychischen Prozesse, die die Einleitung

---

und Aufrechterhaltung zielbezogenen Handelns leisten" zu fassen sind (Ziegler, 1999, S.103). Folgt man diesem Prozessverständnis von Motivation, so ermöglicht das sozialkognitive Motivationsmodell von Dweck und Leggett (1988) eine integrative Betrachtungsweise motivationaler Stärken bzw. Defizite. Die zentrale Idee dieses Modells ist, die Zielorientierung einer Person eingebettet in einen Rahmen selbstbezogener Kognitionen (v.a. der impliziten Persönlichkeitstheorie) zu betrachten. Dweck und Leggett (1988) gehen von mehreren Komponenten aus, die zur Genese von Hilflosigkeit beitragen, d.h. die bei Misserfolg entweder zu meisterndem oder zu hilflosem Verhalten führen.

1. *Zielorientierung*: Die Ziele, die Personen im Leistungskontext verfolgen (*achievement goals*), werden in Lernziele (*learning goals*) und Leistungsziele (*performance goals*) klassifiziert (Elliot, 2005). Lernzielorientierte Personen verfolgen das Ziel, ihre Kompetenzen weiter zu entwickeln. Bei leistungszielorientierten Personen steht die Demonstration der eigenen Kompetenz bzw. eine positive Fähigkeitsbeurteilung im Vordergrund. Mit den jeweiligen Zielorientierungen sind unterschiedliche Muster in Emotion, Kognition und Verhalten verbunden, was vor allem beim Eintritt eines Misserfolges zum Tragen kommt und damit zu bedeutsamen Differenzen führt. Ein Misserfolg führt im Falle von dominierenden Leistungszielen eher zu Hilflosigkeit, da das eigentliche Ziel – eine positive Fähigkeitsbeurteilung – nicht erreicht wurde. Bei dominierender Lernzielorientierung werden in Misserfolgen jedoch auch die Lernchancen und hilfreichen Hinweise für Lernfortschritte gesehen (Dweck, 1999). Derartige Zusammenhänge zeigen sich domainunspezifisch in unterschiedlichen Lernkontexten, u.a. auch bei Studierenden in der Methodenausbildung: So korrelieren Lernziele in Statistik-Kursen positiv mit Selbstwirksamkeit und negativ mit der Prüfungsangst. Leistungsziele hingegen

---

gehen mit einer höheren Prüfungsangst einher (Bandalos, Finney & Geske, 2003).

2. *Implizite Persönlichkeitstheorie (IPT)*: Dweck (1999) unterscheidet zwischen Personen mit stabiler (*entity*) und flexibler (*increment*) Theorie der eigenen Fähigkeiten. Personen mit einer stabilen Persönlichkeitstheorie halten ihre Fähigkeiten für unveränderbare Merkmale. Personen mit einer inkrementellen Persönlichkeitstheorie hingegen glauben, dass ihre Fähigkeiten eine flexible Qualität aufweisen und somit durch Lernen und Anstrengung weiterentwickelt werden können. Personen können domainspezifisch über unterschiedliche Theorien verfügen. Beispielsweise kann ein/e Schüler/in glauben, dass seine/ihre Fähigkeiten in Mathematik unveränderbar sind und gleichzeitig seine/ihre verbalen Fähigkeiten für entwickelbar halten. Beide Theorien (*entity* und *increment*) sind gleich verbreitet, wobei Untersuchungen bei Kindern und Erwachsenen zeigten, dass etwa 40 Prozent zur Entitätstheorie, weitere 40 Prozent zur inkrementellen Theorie tendieren und etwa 20 Prozent unentschlossen sind (Dweck & Molden, 2005).

Die implizite Persönlichkeitstheorie ist für das Lernverhalten und die Kompetenzentwicklung besonders relevant. Empirische Ergebnisse zeigen, dass SchülerInnen mit einer inkrementellen Theorie eher lernzielorientiert sind und Anstrengung als etwas Positives interpretieren, das ihnen hilft neue Kenntnisse zu erwerben. Misserfolge spornen sie zu weiteren Anstrengungen an, da sie Fehler als Lernchancen sehen. Im Unterschied dazu sind SchülerInnen, die an die Stabilität ihrer Fähigkeiten glauben, eher leistungszielorientiert und interpretieren Anstrengung als etwas Negatives, das auf zu geringe Fähigkeiten deutet. Misserfolge führen bei ihnen zu einem Nachlassen der Lernanstrengung.

---

Zwischen den beiden SchülerInnen-Typen konnten auch Leistungsunterschiede nachgewiesen werden: Bei gleichem Vorwissen erreichten die SchülerInnen mit einer flexiblen Theorie bessere Mathematik-Noten als jene SchülerInnen, die ihre Fähigkeiten als unveränderbar ansahen (Blackwell, Dweck & Trzesniewski, 2003). Zusammenfassend führt eine flexible (*increment*) Persönlichkeitstheorie eher zu meisterndem Verhalten, Steigerung der Anstrengung und des Interesses – auch im Fall von Misserfolg – während eine stabile (*entity*) Persönlichkeitstheorie im Falle eines Misserfolgs eher hilfloses Verhalten bewirkt.

3. *Vertrauen in die eigenen Fähigkeiten*: Die erwähnten Zusammenhänge werden jedoch vom Vertrauen in die eigenen Fähigkeiten (*confidence in own abilities*) moderiert (siehe Abbildung 1). So kann die beschriebene hilflose Reaktion durch ein hohes Vertrauen in die eigenen Fähigkeiten zumindest kurzfristig verhindert werden (Dweck, 1999). Ungünstig wirkt sich hingegen die Kombination einer stabilen Fähigkeitstheorie gepaart mit Leistungszielorientierung und geringem Vertrauen in die eigenen Fähigkeiten aus. Durch diese ungünstige Kombination kann die Auseinandersetzung mit einem Lernstoff, wie etwa der Methodenlehre im Psychologiestudium, zu einem "Alptraum" werden (*hilflose Verhaltensreaktion bei Misserfolg*). Andererseits kann der Lernstoff bei Personen mit einer flexiblen Fähigkeitstheorie und/oder einem hohen Vertrauen in die eigenen Fähigkeiten zur positiven Herausforderung werden (*meisterndes Verhalten bei Misserfolg*).

Das sozialkognitive Motivationsmodell (siehe Abbildung 1) erklärt somit systematisch und relativ einfach, wodurch Hilflosigkeit in Lernsituationen bedingt ist. Hilflosigkeit und Angst sind – wie einleitend ausgeführt – bei Psychologiestudierenden in Bezug auf Statistik häufig aufzufinden.

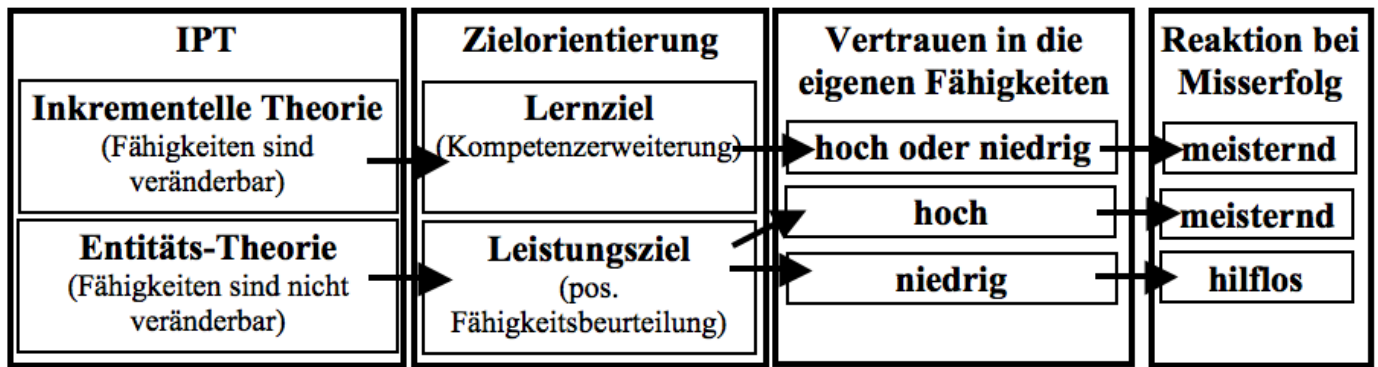


Abbildung 1. Die zentralen Komponenten des sozialkognitiven Motivationsmodells von Dweck und Leggett (1988)

An Hand der bisherigen Ausführungen wurde das Desiderat einer systematischen Analyse der motivationalen Situation von Psychologiestudierenden im Bereich Methodenlehre deutlich. Bisher wurden die problematischen Motivationskomponenten hauptsächlich als Einzelaspekte untersucht. Es fehlt die Analyse dieser Komponenten in einem gemeinsamen Rahmenmodell, die auch als Basis für Interventionsansätze notwendig ist. Das vorgestellte Modell von Dweck und Leggett (1988) bietet zur theoriegeleiteten Systematisierung der motivationalen Befunde eine gute Möglichkeit, da sich für die einzelnen Variablen bereits gezeigt hat, dass sie im Leistungskontext für die Entstehung von Hilflosigkeit und die Erklärung von Leistungsergebnissen relevant sind. In weiterer Folge ist die bisher nicht hinreichend geklärte Frage der Geschlechtsdifferenzen in diesem Kontext ein interessanter Aspekt. Jenseits dieses eher anwendungsbezogenen Fokusses wurde in den bisherigen Ausführungen aber auch evident, dass eine empirische Untersuchung des Dweck'schen Modells in seiner Gesamtheit auf theoretischer Ebene ein klares

---

Forschungsdesiderat darstellt. Vor diesem Hintergrund werden in der vorliegenden Arbeit folgende Fragestellungen untersucht:

1. Wie stellt sich die Leistung und die motivationale Situation von Psychologiestudierenden hinsichtlich der zentralen Variablen des sozialkognitiven Motivationsmodells im Bezug auf die Methodenlehre dar? Zeigen sich in der Motivation oder der Leistung bezogen auf die Methodenlehre Geschlechtsunterschiede?

2. Können die zentralen Annahmen des sozialkognitiven Motivationsmodells von Dweck und Leggett (1988) über die Zusammenhänge der einzelnen Motivationskomponenten bestätigt werden? Können die prognostizierten Verläufe als Profile mittels einer querschnittlichen Typenanalyse identifiziert werden? Unterscheiden sich Studierende mit unterschiedlichen Profilen auch in der Hilfslosigkeit, dem Interesse und dem methodischen Fachwissen?

## 7.4. Methode

### Design, Durchführung und Stichprobe

Die vorliegende Studie ist eine Querschnittserhebung an Studierenden des Diplomstudiums der Psychologie an der Universität Wien. Die Stichprobe umfasst insgesamt 465 Studierende, davon 400 Frauen und 65 Männer. Die Daten wurden in den Studienjahren 2003/04 und 2004/05 erhoben. Dabei wurden im Studienjahr 2003/04 alle Studierenden, die in diesem Studienjahr die Vorlesung "Forschungsmethoden und Evaluation" besuchten, schriftlich befragt und getestet ( $n = 415$ ). Zur Erhöhung der Stichprobe wurden dann im Studienjahr darauf nochmals 50 Personen aus dieser Lehrveranstaltung befragt (die Datenerhebung erfolgte im Rahmen eines größeren Forschungsprojekts und die Ressourcen

---

ließen keine nochmalige Vollerhebung zu). In den untersuchten Variablen zeigten sich keine Unterschiede zwischen den beiden untersuchten Kohorten. Die Studierenden befanden sich im Mittel im achten Semester des Diplomstudiums "Psychologie" ( $M = 8.40$ ;  $SD = 2.43$ ). Das Durchschnittsalter betrug 24.5 Jahre ( $SD = 4.17$  Jahre).

Da es sich bei der Vorlesung um eine für alle Psychologiestudierende verpflichtende Lehrveranstaltung im zweiten Studienabschnitt handelt, kann davon ausgegangen werden, dass es sich dabei um eine repräsentative Stichprobe von Psychologiestudierenden der Universität Wien handelt.

## Erhebungsinstrumente

*Methodisches Fachwissen.* Für das *Fachwissen* in der Methodenlehre wurden zwei Indikatoren erhoben: (1) *Note\_in der Methodenlehre* und (2) *Wissen in der Methodenlehre*. (1) Die *Note in der Methodenlehre* wurde aus dem Diplomprüfungszeugnis des ersten Studienabschnittes (entspricht Vordiplom) entnommen. Bei dieser Note handelt es sich um eine Durchschnittsnote aus vier schriftlichen Vorlesungsprüfungen (VO 1 und 2: Psychologische Methodenlehre und Statistik I und II, VO 3: Testtheorie und Testkonstruktion, VO 4: Qualitative Methoden), Noten aus zwei Übungen zur Psychologischen Methodenlehre und Statistik und der Note aus einem Forschungspraktikum, in dem die Studierenden eine empirische Studie konzipieren und eigenständig durchführen sollen. (2) Das *Wissen in der Methodenlehre* wurde anhand eines Leistungstests zu Methodenlehre operationalisiert, der zu Beginn der Vorlesung "Forschungsmethoden und Evaluation" (1. Kohorte) bzw. zu Beginn der Übung zur Vorlesung (2. Kohorte) vorgegeben wurde (bestehend aus 20 Items;



---

Bewertungsskala: 0-27 Punkte, je höher desto besser; interne Konsistenz:  $\alpha = .53$ ). Die Note in der Methodenlehre bildet daher die Durchschnittsleistung der letzten Semester ab (Notenskala: 1-5, je geringer desto besser), während der Leistungstest ein Maß für das aktuelle Wissen darstellt.

*Stellenwert Methodenlehre.* Die Bedeutung der Methodenlehre für Psychologiestudierende in Relation zu andern Prüfungsfächern wurde über eine offene Frage zum Schwerpunkt im Studium erhoben (*"Mein Studienschwerpunkt innerhalb der Psychologie ist ..."*).

*Motivation.* Die Auswahl der Skalen erfolgte gemäß des sozialkognitiven Motivationsmodells von Dweck und Leggett (1988). Sie wurden domainspezifisch für den Bereich Methodenlehre umformuliert und weisen ein sechsstufiges Antwortformat von (1) "stimme gar nicht zu" bis (6) "stimme völlig zu" auf.

1. Skala *Implizite Persönlichkeitstheorie – Inkrement* in enger Anlehnung an das Instrument von Dweck und Mitarbeitern (Dweck, Chiu & Hong, 1995; Dweck & Henderson, 1988) bestehend aus 4 Items (interne Konsistenz:  $\alpha = .70$ ). Itembeispiel: "Ich bin für Forschungsmethoden nicht sehr begabt – daran kann ich auch nichts ändern." Je höher die Ausprägung ist, desto flexibler ist die implizite Theorie der eigenen Fähigkeiten.
2. Skala *Lernzielorientierung* in Anlehnung an Schober (2002) bestehend aus 4 Items (interne Konsistenz:  $\alpha = .82$ ). Itembeispiel: "Im Bereich Forschungsmethoden geht es mir vor allem darum, meine Fähigkeiten zu steigern."
3. Skala *Leistungszielorientierung* in Anlehnung an Schober (2002) bestehend aus 4 Items (interne Konsistenz:  $\alpha = .81$ ). Itembeispiel: "Im

---

Bereich Forschungsmethoden geht es mir vor allem darum, die Prüfung zu bestehen."

4. Skala *Vertrauen in die eigenen Fähigkeiten* in Anlehnung an Jerusalem und Satow (1999) bestehend aus 4 Items (interne Konsistenz:  $\alpha = .74$ ).  
Itembeispiel: "Es fällt mir leicht, neue Stoffinhalte zu verstehen."
5. Skala *Hilflosigkeit* in Anlehnung an Breitenkopf (1985) bestehend aus 3 Items (interne Konsistenz:  $\alpha = .84$ ). Itembeispiel: "Ich fühle mich in Forschungsmethoden oft überfordert."
6. Skala *Interesse* in Anlehnung an die Skala "Gefühlsbezogene Valenzen" aus dem "Fragebogen zum Studieninteresse" (Schiefele, Krapp, Wild & Winteler, 1993) bestehend aus 4 Items (interne Konsistenz:  $\alpha = .87$ ).  
Itembeispiel "Mir liegt viel daran, im Bereich Forschungsmethoden viel zu wissen."

## 7.5. Ergebnisse

### Fragestellung 1 – Motivationale Situation und fachliche Kompetenz Psychologiestudierender in Bezug auf das Fach Methodenlehre

Bei der Frage nach dem Studienschwerpunkt innerhalb des Psychologiestudiums nennt lediglich ein Prozent der Studierenden ( $n = 4$ ) die Methodenlehre (siehe Tabelle 1). Führend ist die "Klinische Psychologie", die knapp 40 Prozent der befragten Studierenden als ihren Studienschwerpunkt angaben. Am zweiten Platz steht die Gruppe jener, die (noch) keinen konkreten Schwerpunkt gewählt haben. Dazu zählt etwa ein Viertel der Studierenden. Im Mittelfeld liegen die Fächer "Wirtschaftspsychologie" (12.5%), "Sozialpsychologie" (9.5%) und "Entwicklungspsychologie" (8.6%). Die Analyse

der Geschlechterverteilung in der Wahl der Studienschwerpunkte ergab einen signifikanten Unterschied [ $\chi^2$  (10,  $N = 463$ ) = 23.72,  $p < .01$ ], der – wie die Betrachtung der Kreuztabelle nahelegt – im Wesentlichen durch die unterschiedliche Präferenz bzgl. Klinischer Psychologie bedingt ist. Die absoluten Angaben im Bereich Methodenlehre sind zu gering um Geschlechtsunterschiede interpretieren zu können.

*Tabelle 1: Studienschwerpunkte innerhalb der Psychologie*

Schwerpunkt	Häufigkeit			Prozent		
	Gesamt	Frauen	Männer	Gesamt	Frauen	Männer
Klinische Psychologie	171	153	18	36.8	38.3	27.7
kein definierter Schwerpunkt	112	94	18	24.1	23.5	27.7
Wirtschaftspsychologie	58	48	10	12.5	12.0	15.4
Sozialpsychologie	44	37	7	9.5	9.3	10.8
Entwicklungspsychologie	40	37	3	8.6	9.3	4.6
Diagnostik	15	13	2	3.2	3.3	3.1
Bildungspsychologie	13	10	3	2.8	2.5	4.6
Kriminologie	6	5	1	1.3	1.3	1.5
Methodenlehre	4	1	3	0.9	0.3	4.6
Allgemeine Psychologie	1	1	0	0.2	0.3	0.0
Biologische Psychologie	1	1	0	0.2	0.3	0.0
Gesamt	465	400	65	100	100	100

Anmerkung. Da die Studierenden nach ihrem inhaltlichen Studienschwerpunkt gefragt wurden, entsprechen nicht alle Schwerpunkte einschlägigen Studienfächern. Angegeben sind die Häufigkeiten und Prozentwerte, mit denen bestimmte Bereiche der Psychologie von den Psychologiestudierenden als Studienschwerpunkt genannt wurden. Im Fach Allgemeine Psychologie fand zum Zeitpunkt der Erhebung eine Neubesetzung statt. Für Studierende war die Allgemeine Psychologie daher nicht klar als Fach erkennbar.

Genauere und systematische Informationen über die aktuelle Motivation der Studierenden in Bezug auf die Methodenlehre enthalten die ausgewerteten Fragebogenskalen. Im Gegensatz zu den Literaturbefunden ist die *motivationale Situation* der Studierenden im Fach Methodenlehre in der vorliegenden Stichprobe überraschend positiv. Betrachtet man den Durchschnitt der einzelnen motivationalen Determinanten, so liegen bis auf die Leistungszielorientierung alle Mittelwerte auf der positiven Seite der jeweiligen Skala. Die Skalen Hilfflosigkeit und Interesse befinden sich allerdings im mittleren Skalenbereich und sind somit

---

von den jeweiligen Idealwerten – niedrige Hilflosigkeit sowie hohes Interesse – weit entfernt (siehe Tabelle 2). Zudem weist die hohe Varianz sowie die Verteilung bei der Hilflosigkeit darauf hin, dass ein gewisser Prozentsatz Studierender von einem sehr hohen Hilflosigkeitserleben berichtet. Tendenziell betrifft dies eher weibliche Studierende, da bestehende geschlechtsspezifische Unterschiede in den motivationalen Merkmalen ausschließlich zugunsten der männlichen Studierenden bestehen (siehe Tabelle 2): Männer weisen in Bezug auf die Methodenlehre eine flexiblere implizite Persönlichkeitstheorie, eine geringere Hilflosigkeit und ein höheres Vertrauen in die eigenen Fähigkeiten auf als Frauen, wobei die praktische Relevanz dieser Unterschiede als niedrig einzustufen ist (siehe Effektstärken in Tabelle 2). Keine Unterschiede bestehen hingegen in der Lern- und Leistungszielorientierung sowie dem Interesse.

Die Betrachtung des methodischen Fachwissens auf der Basis der Note aus dem Diplomprüfungszeugnis des ersten Studienabschnittes (*Note in der Methodenlehre*) ergab einen Durchschnittswert von 2.4. Das entspricht auf der fünfstufigen Notenskala einer Leistung zwischen "gut" und "befriedigend". Geschlechtsunterschiede konnten nicht festgestellt werden. Jedoch erreichten die Männer im aktuellen Wissenstest (*Wissen in der Methodenlehre*) eine höhere Punkteanzahl als die Frauen, wobei dieser Unterschied einer kleinen Effektstärke entspricht (siehe Tabelle 2). Die mittlere Leistung im Wissenstest lag knapp über der Hälfte der maximal erreichbaren Punkteanzahl. In Anbetracht der differentiellen Geschlechtsunterschiede wurden getrennt für Männer und Frauen Korrelationen zwischen den beiden Indikatoren (Note und Wissenstest) gerechnet. Dabei wurde für beide Geschlechter ein bedeutsamer, jedoch inhaltlich geringer Zusammenhang festgestellt, was darauf deutet, dass die Studierenden entweder den Lernstoff schnell vergaßen, oder dass der Note und

dem Leistungstest andere Beurteilungsmaßstäbe zu Grunde lagen. Diesbezüglich gab es jedoch keine Geschlechtsunterschiede, da in der Höhe der Korrelation sich kein statistisch bedeutsamer Unterschied zwischen Männern und Frauen zeigte ( $r_{\text{Männer}} = -.31$ ;  $r_{\text{Frauen}} = -.22$ ).

*Tabelle 2: Geschlechtsunterschiede in Fachkompetenz und Motivation hinsichtlich Methodenlehre*

		Gesamt	Frauen	Männer	<i>F</i>	<i>df</i>	<i>p</i>	$\eta^2$
Wissen in Methodenlehre	M	14.45	14.28	15.51	6.87	1;439	.009	0.02
	SD	3.54	3.48	3.69				
Note in Methodenlehre	M	2.43	2.44	2.42	0.03	1;439	.839	
	SD	0.61	0.60	0.64				
IPT Inkrement	M	4.76	4.72	5.00	5.66	1;439	.016	0.01
	SD	0.84	0.86	0.67				
Lernzielorientierung	M	4.47	4.48	4.39	1.27	1;439	.288	
	SD	0.76	0.74	0.92				
Leistungszielorientierung	M	3.61	3.62	3.57	0.10	1;439	.752	
	SD	1.06	1.05	1.09				
Vertrauen in die eigenen Fähigkeiten	M	4.58	4.55	4.75	3.65	1;439	.052	0.01
	SD	0.75	0.75	0.71				
Hilflosigkeit	M	2.86	2.92	2.46	9.26	1;439	.002	0.02
	SD	1.13	1.11	1.15				
Interesse	M	3.70	3.68	3.81	0.76	1;439	.367	
	SD	0.98	0.95	1.13				

*Anmerkung.* Angegeben sind Mittelwerte (*M*) und Standardabweichungen (*SD*) von Gesamtstichprobe, Frauen und Männer; Ergebnisse der multivariaten Varianzanalyse mit dem unabhängigen Faktor "Geschlecht"; Effektstärken ( $\eta^2$ ); Gesamtmodell-Geschlecht:  $F(8;431) = 2.37$ ,  $p = .016$ ,  $\eta^2 = 0.04$ . Der Skalenrange reicht im Wissen in Methodenlehre von 0-27, in der Note Methodenlehre von 1-5 und in den restlichen Skalen von 1-6.

## Fragestellung 2 – Analyse der zentralen Annahmen des sozialkognitiven Motivationsmodell von Dweck und Leggett (1988)

Im Folgenden wurde der Frage nachgegangen, ob die zentralen Annahmen des sozialkognitiven Motivationsmodells von Dweck und Leggett (1988) bestätigt werden können. Dazu sollten die im Modell postulierten Verläufe mittels Typenanalyse als Profile in der Studierendenstichprobe identifiziert werden. Die

---

zentralen Variablen des Modells sind die implizite Persönlichkeitstheorie, Lern- und Leistungszielorientierung sowie das Vertrauen in die eigenen Fähigkeiten. Durch den Einsatz von Clusteranalysen sollten (möglichst theoriekonforme) Profile (= Typen) identifiziert werden. Unterschiede zwischen den Typen in den abhängigen Variablen (Hilflosigkeit, Interesse und methodisches Fachwissen) in erwarteter Richtung wären weitere Hinweis für die Gültigkeit des Modells.

Die Typenbildung erfolgte mittels eines dreistufigen clusteranalytischen Verfahrens. Im ersten Schritt wurden Cluster nach der single linkage Methode gebildet, um mögliche Ausreißer zu identifizieren (Backhaus, Erichson, Plinke & Weiber, 2000). Es wurden vier Ausreißer identifiziert, die aus der weiteren Analyse ausgeschlossen wurden. Im zweiten Schritt wurde eine hierarchische Clusteranalyse nach dem Ward-Verfahren durchgeführt. Als Distanzmaß wurde die quadrierte euklidische Distanz verwendet. Auf Basis einer 3-Clusterlösung, die im dritten Schritt mit Hilfe der k-means Methode optimiert wurde, lassen sich drei in etwa gleich häufig besetzte Motivationstypen beschreiben. Hinsichtlich der Güte der Clusteranalyse ist anzuführen, dass die Cluster im Vergleich zur Gesamtstichprobe sehr homogen sind;  $\Lambda = 0.160$  (Aldenderfer, 1982). In der Verteilung von Männern und Frauen auf die drei Typen gibt es keine Unterschiede [ $\chi^2(2, N = 461) = 2.89, p > .05$ ]. Die Geschlechtsunterschiede auf Variablenebene (in der impliziten Persönlichkeitstheorie und tendenziell im Vertrauen in die eigenen Fähigkeiten) zeigen sich also nicht in den personenbezogenen Analysen der Motivationstypen. Dies entspricht durchaus dem sozialkognitiven Modell, da die beiden Aspekte als voneinander unabhängig postuliert werden. Die Bezeichnung der Typen erfolgte nach den Kernmerkmalen für die jeweiligen Typen.

- Typ 1, die "*Kompetenzorientierten Studierenden*" ( $n = 145$ ) lernen vor allem, um Kompetenzen zu erwerben und sind kaum leistungszielorientiert. Zusätzlich haben sie ein hohes Vertrauen in ihre eigenen Fähigkeiten und sehen diese als veränderbar an.
- Typ 2, die "*Leistungsorientierten Studierenden*" ( $n = 160$ ) sind den "*Kompetenzorientierten Studierenden*" relativ ähnlich. Der wesentliche Unterschied liegt im Ausmaß der Leistungszielorientierung. Die "*Leistungsorientierten Studierenden*" lernen nicht nur, um Kompetenz zu erwerben, ihnen sind gute Noten und die Demonstration des Könnens ebenso wichtig. Neben der hohen Lernzielorientierung haben sie also auch eine hohe Leistungszielorientierung.

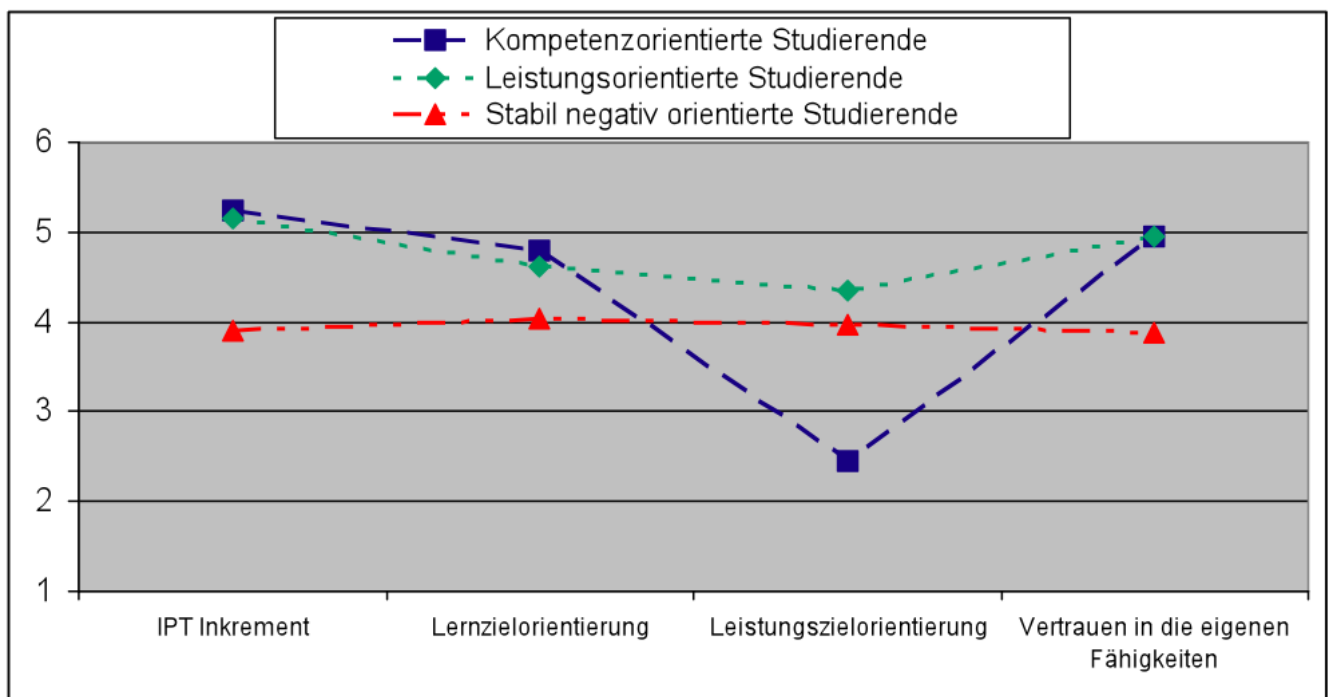


Abbildung 2. Motivationstypen im Sinne des sozialkognitiven Motivationsmodells.

- Typ 3, die "*Stabil negativ orientierten Studierenden*" ( $n = 156$ ) sind durch die geringere Überzeugung in die Veränderbarkeit der eigenen Fähigkeiten

---

gekennzeichnet. Gleichzeitig haben diese Studierenden auch ein geringeres Vertrauen in die eigenen Fähigkeiten. Zusätzlich zu diesen beiden ungünstigeren Lernvoraussetzungen sind die "Stabil negativ orientierten Studierenden" im Vergleich zu den anderen Typen auch weniger lernzielorientiert.

Gemäß dem sozialkognitiven Motivationsmodell würden Typ 1 und 2 nach einem Misserfolg meisternd reagieren, da Personen beider Typen über eine flexible implizite Persönlichkeitstheorie verfügen, lernzielorientiert sind und ein hohes Vertrauen in die eigenen Fähigkeiten besitzen. Typ 3 weist in diesen drei Variablen geringere Werte auf, ist aber dafür stärker leistungszielorientiert. Bei diesem Typ wäre eine hilflose Reaktion bei Misserfolg zu erwarten. Zur Klärung dieser Frage sowie allfälliger Leistungsunterschiede wurde eine multivariate Varianzanalyse (GLM mit *Wissen in der Methodenlehre*, *Note in der Methodenlehre*, *Interesse* und *Hilflosigkeit* als abhängige Variablen und den *Motivationstypen* als unabhängiger Faktor) durchgeführt. In der Varianzanalyse zeigte sich ein multivariater Effekt [ $F(8, 868) = 27.16, p < .001, \eta^2 = .20$ ], der sich auf univariater Ebene durch Unterschiede im Wissen Methodenlehre [ $F(2, 436) = 11.46, p < .001, \eta^2 = .05$ ], der Note Methodenlehre [ $F(2, 436) = 8.81, p < .001, \eta^2 = .04$ ], im Interesse [ $F(2, 436) = 64.90, p < .001, \eta^2 = .23$ ] und in der Hilflosigkeit [ $F(2, 436) = 116.56, p < .001, \eta^2 = .35$ ] manifestierte. Post-hoc-Tests (Scheffé) bestätigten, dass sich die "Stabil negativ orientierten Studierenden" (Typ 3) in allen Variablen zu ihrem Nachteil signifikant von den beiden anderen Typen unterscheiden. Da sich die Typen 1 und 2 in keiner der Variablen unterscheiden, wurden diese beiden Typen als "meisternde Studierende" zusammengefasst und mit Typ 3 verglichen. In der Varianzanalyse zeigte sich ein multivariater Effekt [ $F(4, 434) = 68.66, p < .001, \eta^2 = 0.39$ ], der



sich auf univariater Ebene durch Unterschiede im Wissen in der Methodenlehre [ $F(1, 437) = 22.24, p < .001, \eta^2 = 0.05$ ], der Note in der Methodenlehre [ $F(1, 437) = 15.86, p < .001, \eta^2 = 0.04$ ], im Interesse [ $F(1, 437) = 123.46, p < .001, \eta^2 = 0.22$ ] und in der Hilfflosigkeit [ $F(1, 437) = 233.15, p < .001, \eta^2 = 0.35$ ] manifestierte. Die "meisternden Studierenden" zeigten in all diesen Variablen, vor allem aber in Interesse und Hilfflosigkeit, wesentlich bessere Werte als die "Stabil negativ orientierten Studierenden".

*Tabelle 3: Unterschiede zwischen den drei Motivationstypen*

		Typ 1 Kompetenz- orientierte Studierende	Typ 2 Leistungs- orientierte Studierende	Typ 3 Stabil negativ orientierte Studierende	Gesamt
<i>n</i>		145	160	156	461
Wissen in Methodenlehre <sup>a,b</sup>	M	15.16	14.74	13.44	14.43
	SD	3.76	3.34	3.29	3.53
Note in Methodenlehre <sup>a,b</sup>	M	2.30	2.39	2.59	2.43
	SD	0.65	0.54	0.60	0.61
Hilfflosigkeit <sup>a,b</sup>	M	2.35	2.41	3.79	2.86
	SD	0.92	0.89	0.90	1.12
Interesse <sup>a,b</sup>	M	4.15	3.93	3.06	3.70
	SD	0.91	0.89	0.78	0.98

*Anmerkung.* Angegeben sind die Mittelwerte (*M*) und Standardabweichungen (*SD*) der drei Motivationstypen und der Gesamtstichprobe sowie die Personenzahl jedes Clusters (*n*). Dargestellt sind jene Skalen, die nicht zur Typenbildung herangezogen wurden.

Ergebnisse des Scheffé post-hoc Tests:

<sup>a</sup> sign. Unterschied ( $p < 0.05^*$ ) zwischen Typ 1 und Typ 3.

<sup>b</sup> sign. Unterschied ( $p < 0.05^*$ ) zwischen Typ 2 und Typ 3.

## 7.6. Diskussion

Intention dieser Studie war es, einerseits Motivationsprobleme von Psychologiestudierenden im Bereich Methodenlehre systematisch im Rahmen eines etablierten Motivationsmodells zu betrachten und andererseits zentrale Annahmen des sozialkognitiven Motivationsmodells (Dweck & Leggett, 1988) zu

---

prüfen. Im Folgenden sollen die wichtigsten Befunde zusammengefasst und sich daraus ergebende theoretische sowie praktische Schlussfolgerungen diskutiert werden.

Vergleicht man die Ergebnisse der vorliegenden Studie mit bisherigen Befunden in diesem Bereich, so sind die hier berichteten Motivationsprobleme der Psychologiestudierenden bei weitem nicht so dramatisch, wie in der bisherigen Literatur dargestellt (Birenbaum & Eylath, 1994; Onwuegbuzie, 2004; Onwuegbuzie & Wilson, 2003; Stark & Mandl, 2005; Zeidner, 1991). Trotz dieser verhältnismäßig positiven motivationalen Situation der Psychologiestudierenden liegt die Methodenlehre nicht im Hauptinteresse der Studierenden (nur ein Prozent der Studierenden wählt Methodenlehre als Studienschwerpunkt). Bezogen auf die absoluten Werte (Einschätzungen auf Motivationsskalen) geben die Studierenden in Bezug auf die Methodenlehre ein relativ hohes Vertrauen in die eigenen Fähigkeiten an, sehen diese Fähigkeiten eher als veränderbar an und stufen sich als lernzielorientiert ein. Weniger positiv ausgeprägt sind das Interesse an der Methodenlehre sowie die Werte hinsichtlich der Hilflosigkeit. Dieser scheinbare Widerspruch zum postulierten Modell von Dweck und Leggett (1988) zeigt sich allerdings nur bei variablenorientierter Betrachtung. Aus personenorientierter Perspektive ist dieses widersprüchliche Muster bei den Studierenden nicht zu finden. Hier entsprechen die drei generierten Typen von Studierenden dem theoretischen Modell von Dweck und Leggett (1988).

Hinsichtlich Geschlechtsunterschiede in der motivationalen Situation der Studierenden wurde die Mehrheit der bisherigen Befunde bestätigt (Benson, 1989; Onwuegbuzie & Wilson, 2003; Roberts & Saxe, 1982). So erachten Frauen ihre Fähigkeiten in Bezug auf die Methodenlehre weniger veränderbar als Männer

---

und erleben sich als hilfloser im Bereich Methodenlehre. Darüber hinaus weisen weibliche Studierende in dieser Studie ein tendenziell geringeres Vertrauen in die eigenen Fähigkeiten als männliche auf. Bei Berücksichtigung der Erkenntnis, dass diese Merkmale wichtige Determinanten des Leistungsverhaltens darstellen (Dweck & Molden, 2005; Multon, Brown & Lent, 1991; Schiefele, Krapp & Schreyer, 1993), muss die motivationale Situation der weiblichen Psychologiestudierenden als eher problematisch eingestuft werden. Nachvollziehbar wird dieser Befund angesichts der oft beschriebenen ungünstigeren motivationalen Voraussetzungen von Mädchen und Frauen im mathematisch-naturwissenschaftlichen Bereich (vgl. Schober, 2002).

Auch im generell eher niedrigen methodischen Fachwissen zeigen sich Geschlechtsunterschiede mit mittlerer Effektstärke zu Ungunsten der Frauen. Die Psychologiestudentinnen erzielten im Mittel nur knapp die Hälfte der maximal erreichbaren Punkteanzahl eines vorgegebenen Wissenstests, was nach vielfach gebräuchlichen Beurteilungskriterien gerade noch der Note 4, also "genügend" oder "ausreichend" entsprechen würde. Im Gegensatz dazu erreichten Männer hierbei bessere Resultate als Frauen. Diese Leistungsunterschiede könnten auch Ausdruck eines "underachievement" von Frauen sein, wonach sie aufgrund der männlich konnotierten Statistik-Lehre schwächere Leistungen zeigen, als es nach ihrer Leistungsfähigkeit zu erwarten wäre (Hannover & Rau, in Druck).

Andererseits zeigt Schram (1996) in ihrer Metaanalyse, dass Geschlechtsunterschiede im methodischen Fachwissen zu Gunsten der Männer v.a. in jenen Studien gefunden werden, bei denen die Ergebnisse auf (Leistungs)tests beruhen. Werden bei dem Vergleich Noten herangezogen, bei denen auch die Anwesenheit (Mitarbeit) und Hausübungen berücksichtigt werden, zeigt sich ein gegenteiliges Bild. Zwar kann dieser Befund nicht zur

---

Gänge bestätigt werden, doch zeigen sich in den vorliegenden Daten auf der Ebene der Noten (Durchschnittsnote der Lehrveranstaltungen im Fach Methodenlehre vom ersten Studienabschnitt) zumindest keine Geschlechtsunterschiede. Diese scheinbar gegensätzlichen Befunde werden durch die geringe Korrelation zwischen Note und den Ergebnissen des Leistungstests verständlich. Ursache dieser geringen Korrelation könnten unterschiedliche Beurteilungsmaßstäbe sein. Dafür sprechen auch einige internationale Befunde, die aufzeigen, dass sich die Nivellierung der Geschlechtsunterschiede in den Noten nicht gleichermaßen in Leistungstests wieder findet (vgl. hierzu auch Kenney-Benson, Pomerantz, Ryan & Patrick, 2006). So könnte in die Notenvergabe eher Gewissenhaftigkeit und Mitlernen mit einfließen und von Vorteil sein, während bei Leistungstests ausschließlich der Umgang mit neuen, unbekannten Aufgaben und Problemlösestrategien geprüft wird (vgl. Schram, 1996).

Das sozialkognitive Motivationsmodell (Dweck & Leggett, 1988) wird durch die Daten insofern gestützt, als mittels Clusteranalyse drei Typen gefunden wurden, welche die postulierten Eigenheiten aufweisen. So zeigen Studierende mit einer stabileren impliziten Persönlichkeitstheorie, geringeren Lernzielorientierung und einem geringerem Vertrauen in die eigenen Fähigkeiten (*"Stabil negativ orientierte Studierende"*) auch eine höhere Hilflosigkeit und geringeres Fachwissen als die anderen beiden Typen (*"Kompetenzorientierte Studierende"* und *"Leistungsorientierte Studierende"*), die ein höheres Interesse an der Methodenlehre aufweisen. Die großen Effektstärken in Bezug auf Interesse und Hilflosigkeit deuten auf die hohe Relevanz dieser Unterschiede hin. Trotz der Ähnlichkeit zwischen den beiden letzteren Typen, weisen sie Unterschiede auf, die theoretisch bedeutsam sind. Einerseits belegen diese

---

Unterschiede, dass Lern- und Leistungszielorientierung nicht zwei Ausprägungen einer Dimension sind, sondern durchaus, wie von einigen AutorInnen vertreten (Elliot, 2005), als zwei unabhängige Dimensionen betrachtet werden sollten. Dies könnte durch eine Erweiterung des Modells von Dweck und Leggett (1988) berücksichtigt werden. Andererseits legen die Befunde bzgl. Leistung die Interpretation nahe, dass bei einer hohen Lernzielorientierung die Ausprägung der Leistungszielorientierung keine Auswirkung auf die Leistung hat. Auch dieser Punkt wird in der Literatur umstritten diskutiert (Elliot, 2005).

In Anbetracht der Stichprobe (nur Universität Wien, nur Daten aus der Psychologie) stellt sich natürlich die Frage der Generalisierbarkeit der Ergebnisse. Dies betrifft einerseits die Generalisierbarkeit auf andere sozialwissenschaftliche Studiengänge und andererseits auf andere Studienstandorte. Bezug nehmend auf Diehl (1993), der motivationale Probleme in der Methodenausbildung nicht nur auf das Psychologiestudium beschränkt sieht, scheinen die Ergebnisse mit entsprechender Vorsicht jedoch durchaus auch auf andere sozialwissenschaftliche Studiengänge übertragbar zu sein. Analog den Untersuchungen im schulischen Bereich darf allerdings die Systemkomponente nicht völlig außer Acht gelassen werden, womit natürlich sowohl die Generalisierbarkeit auf Ebene von Studienfächern als auch auf Ebene von Studienstandorten eingeschränkt ist. Trotz dieser Einschränkung lassen sich jedoch mit Blick auf bisherige Befunde in diesem Bereich einige interessante Aspekte diskutieren.

Die Ergebnisse der Studie liefern sowohl theoretische als auch praktische Implikationen. Theoretisch bietet die Studie Hinweise auf die Gültigkeit des sozialkognitiven Motivationsmodells in seiner Gesamtheit. Bisher untersuchten viele Studien Teile des Modells (vgl. z.B. Heller, Finsterwald & Ziegler, 2001),

---

beispielsweise den positiven Zusammenhang zwischen einer flexiblen impliziten Persönlichkeitstheorie und der Leistung (Henderson & Dweck, 1990), nur wenige aber das Modell in seiner Gesamtheit (z.B. Schober, 2001). Die vorliegende Studie leistet diesbezüglich den ersten Schritt, weitere Analysen v.a.

Pfadanalysen sind jedenfalls noch zu machen. Weiters kann auf Grund der Studie eine Erweiterung des Modells unter Berücksichtigung der von anderen auch angenommenen Zweidimensionalität von Lern- und Leistungszielorientierung empfohlen werden.

Praktisch bietet diese Studie konkrete Ansatzpunkte für die Förderung der Motivation von Psychologiestudierenden bezogen auf die Methodenlehre. Auch wenn die Motivationswerte im Mittel bei weitem nicht so dramatisch sind, wie aufgrund früherer Studien zu erwarten war, sind sie dennoch bei rund einem Drittel der Studierenden (den "*Stabil negativ orientierten Studierenden*") als problematisch zu bezeichnen. Als Konsequenz der Studienergebnisse müssten v.a. drei Aspekte gefördert werden: (1) eine flexible implizite Persönlichkeitstheorie, (2) eine stärkere Lernzielorientierung sowie (3) ein höheres Vertrauen in die eigenen Fähigkeiten. Einerseits sind diese Konstrukte keine stabilen Persönlichkeitseigenschaften und können durch Interventionen positiv beeinflusst werden, andererseits führen Verbesserungen in diesen wenigen Bereichen zu einer weitreichenden Verbesserung des Lernens und der Leistung (Dweck & Molden, 2005). Im Bereich der Methodenausbildung liegen bereits erfolgreiche Ansätze vor (Gräsel & Mandl, 1999; Schober, Wagner, Reimann, Atria & Spiel, 2006; Ziegler & Pöppel, 1995), die neben der Wissensvermittlung auch die erwähnten motivationalen Konstrukte systematisch (explizit oder auch nur implizit) fördern. Eine Schlüsselrolle spielt hier beispielsweise konstruktives und individuelles Leistungsfeedback, das sich auf alle drei erwähnten Aspekte positiv

---

auswirken kann. Neue Möglichkeiten der Motivationsförderung v.a. bei stark überlaufenen Lehrveranstaltungen bietet auch der blended learning Ansatz, wie er beispielsweise in dem Projekt "Vienna E-lecturing" angewendet wird (Schober, Wagner, Reimann & Spiel, 2008).

---

## 7.7. Literaturverzeichnis

- Aldenderfer, M.S. (1982). Methods of Cluster Validation for Archaeology. *World Archaeology*, 14, 61-72.
- Backhaus, K., Erichson, B., Plinke, W. & Weiber, R. (2000). *Multivariate Analysemethoden. Eine anwendungsorientierte Einführung*. Berlin: Springer.
- Bandalos, D.L., Finney, S.J. & Geske, J.A. (2003). A model of statistics performance based on achievement goal theory. *Journal of Educational Psychology*, 95, 604-616.
- Benson, J. (1989). Structural components of statistical test anxiety in adults: An exploratory model. *Journal of Experimental Education*, 57, 247-261.
- Birenbaum, M. & Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational sciences. *Educational Research*, 36, 93-98.
- Blackwell, L., Dweck, C.S. & Trzesniewski, K. (2003). Implicit theories of intelligence predict achievement across and adolescent transition: A longitudinal study and an intervention. Unpublished manuscript, Columbia University at New York.
- Blalock, H.M., Jr. (1987). Some general goals in teaching statistics. *Teaching Sociology*, 15, 164-172.
- Breitkopf, L. (1985). Die Hilfslosigkeitsskala. *Diagnostica*, 31, 221-233.
- Cherian, V. & Glencross, M. (1997). Sex, socioeconomic status, and attitude toward statistics among postgraduate education students. *Psychological Reports*, 80, 1385-1386.
- Diehl, J.M. (1993). Statistik-Ausbildung im Psychologie-Grundstudium. In G. Krampen & H. Zayer (Hrsg.), *Psychologische Aus-, Fort- und Weiterbildung in den alten und neuen Ländern* (S. 18-29). Bonn: Dt. Psychologen-Verlag.
- Dresel, M., Stöger, H. & Ziegler, A. (2005). Schul- und Klasseneffekte bei motivationalen Variablen. Ulmer Forschungsberichte aus der Pädagogischen Psychologie Nr. 12. Ulm: Universität, Abt. Pädagogische Psychologie.
- Dweck, C.S. (1999). *Self-theories: Their role in motivation, personality, and development*. Philadelphia: Psychology-Press.
- Dweck, C.S., Chiu, C. & Hong, Y. (1995). Implicit theories and their role in judgements and reactions: A world from two perspectives. *Psychological Inquiry*, 6, 267-285.



- 
- Dweck, C.S. & Henderson, V.L. (1988). Theories of intelligence: Background and measures. Unpublished manuscript, University of Illinois.
- Dweck, C.S. & Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.
- Dweck, C.S. & Molden, D.C. (2005). Self theories: Their impact on competence, motivation, and acquisition. In A.J. Elliot & C.S. Dweck (Hrsg.), *Handbook of competence and motivation*. New York: The Guilford Press.
- Elliot, A.J. (2005). A conceptual history of the achievement goal construct. In A.J. Elliot & C.S. Dweck (Hrsg.), *Handbook of competence and motivation*. New York: The Guilford Press.
- Engeser S., Rheinberg, F., Vollmeyer, R. & Bischoff, J. (2005). Motivation, Flow-Erleben und Lernleistung in universitären Lernsettings. *Zeitschrift für Pädagogische Psychologie*, 19, 159-172.
- Gal, I. & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education*, 2, 1-54.
- Giesler, J.M. (1998). Analysen zur Veränderung fachbezogener Interessen und Einstellungen von Studierenden des Fachs Psychologie in den beiden ersten Studiensemestern. In G. Krampen & H. Zayer (Hrsg.), *Psychologiedidaktik und Evaluation I. Konzepte, Erfahrungsberichte und empirische Untersuchungsbefunde aus Anwendungsfeldern der Aus-, Fort- und Weiterbildung* (S. 27-41). Bonn: Deutscher Psychologen Verlag.
- Gräsel, C. & Mandl, H. (1999). Problemorientiertes Lernen in der Methodenausbildung des Pädagogikstudiums (Forschungsbericht No. 111). München: Universität, Institut für Pädagogische Psychologie und Empirische Pädagogik.
- Gruber, H. & Renkl, A. (1996). Alpträume sozialwissenschaftlicher Studierender: Empirische Methoden und Statistik. In J. Lompscher & H. Mandl (Hrsg.), *Lehr- und Lernprobleme im Studium. Bedingungen und Veränderungsmöglichkeiten* (S. 118-130). Bern: Huber.
- Hannover, B. & Rau, M. (in Druck). Geschlechtergerechte Instruktion am Beispiel der Statistik-Lehre im Psychologiestudium. In C. Spiel, R. Reimann, B. Schober & P. Wagner (Hrsg.), *Bildungspsychologie*. Göttingen: Hogrefe.

- 
- Heller, K.A., Finsterwald, M. & Ziegler, A. (2001). Implicit theories of german mathematics and physics teachers on gender specific giftedness and motivation. *Psychologische Beiträge*, 43, 172-189.
- Henderson, V.L. & Dweck, C.S. (1990). Achievement and motivation in adolescence: A new model and data. In S. Feldman & G. Elliott (Hrsg.), *At the threshold: The developing adolescent* (S. 308-329). Cambridge, MA: Harvard University Press.
- Hilton, S.C., Schau, C. & Olsen, J.A. (2004). Survey of attitudes toward statistics: Factor structure invariance by gender and by administration time. *Structural Equation Modelling*, 11, 92-109.
- Jerusalem, M. & Satow, L. (1999). Schulbezogene Selbstwirksamkeitserwartung (WIRKSCHUL). In R. Schwarzer & M. Jerusalem (Hrsg.), *Skalen zur Erfassung von Lehrer- und Schülermerkmalen*. Berlin: FU.
- Kenney-Benson, G.A., Pomerantz, E.M., Ryan, A.M. & Patrick, H. (2006). Sex Differences in Math Performance: The Role of Children's Approach to Schoolwork. *Developmental Psychology*, 42, 11-26.
- Multon, K., Brown, S. & Lent, R. (1991). Relation of self-efficacy beliefs to academic outcomes: a meta-analytic investigation. *Journal of Counseling Psychology*, 38, 30-38.
- Onwuegbuzie, A.J. (2000). Statistics anxiety and the role of self-perceptions. *The Journal of Educational Research*, 93, 323-330.
- Onwuegbuzie, A.J. (2003). Modeling statistics achievement among graduate students. *Educational and Psychological Measurement*, 63, 1020-1038.
- Onwuegbuzie, A.J. (2004). Academic procrastination and statistics anxiety. *Assessment and Evaluation in Higher Education*, 29, 3-19.
- Onwuegbuzie, A.J. & Wilson, V.A. (2003). Statistics anxiety: Nature, etiology, antecedents, effects, and treatments - a comprehensive review of the literature. *Teaching in Higher Education*, 8, 195-209.
- Pintrich, P.R. & Schunk, D.H. (1996). *Motivation in education: Theory, research and application*. Englewood Cliffs, NJ: Simon & Schuster.
- Roberts, D.M. & Saxe, J.E. (1982). Validity of a statistics attitude survey: A follow-up study. *Educational and Psychological Measurement*, 42, 907-912.

- 
- Schiefele, U., Krapp, A. & Schreyer, I. (1993). Metaanalyse des Zusammenhangs von Interesse und schulischer Leistung. *Zeitschrift für Entwicklungspsychologie und Pädagogische Psychologie*, 25, 120-148.
- Schiefele, U., Krapp, A., Wild, K.-P. & Winteler, A. (1993). Der "Fragebogen zum Studieninteresse" (FSI). *Diagnostica*, 39, 335-351.
- Schober, B. (2001). Implicit personality theories about the stability of behaviour and aspects of volitional behavior control - necessary expansions of Carol Dweck's Motivation Process Model? *Psychologische Beiträge*, 43, 77-99.
- Schober, B. (2002). *Entwicklung und Evaluation des Münchner Motivationstrainings (MMT)* (Theorie und Forschung, Bd. 763). Regensburg: Roderer.
- Schober, B., Wagner, P., Reimann, R., Atria, M. & Spiel, C. (2006). Teaching Research Methods in an Internet-Based Blended-Learning Setting. *Methodology*, 2, 73-82.
- Schober, B., Wagner, P., Reimann, R. & Spiel, C. (2008). Vienna E-Lecturing (VEL): Learning how to learn self-regulated in an internet based blended learning setting. *International Journal on E-Learning*, 7, 703-723.
- Schram, C.M. (1996). A Meta-Analysis of gender differences in applied statistics achievement. *Journal of Educational and Behavioral Statistics*, 21, 55-70.
- Stark, R. & Mandl, H. (2000). Training in empirical research methods: Analysis of problems and intervention from a motivational perspective. In J. Heckhausen (Hrsg.), *Motivational Psychology of Human Development*. Amsterdam: Elsevier.
- Stark, R. & Mandl, H. (2005). Lernen mit einer netzbasierten Lernumgebung im Bereich empirischer Forschungsmethoden. *Unterrichtswissenschaft*, 33, 3-29.
- Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology*, 61, 319-328.
- Ziegler, A. (1999). Motivation. In C. Perleth & A. Ziegler (Hrsg.), *Pädagogische Psychologie* (S. 103-113). Bern: Hans Huber.
- Ziegler, A. & Pöppel, J. (1995). Reattributionstraining mit Statistikstudentinnen. In R. Arbinger & R.S. Jäger (Hrsg.), *Zukunftsperspektiven empirisch-pädagogischer Forschung* (S. 280-291). Landau: Verlag Empirische Pädagogik.

---

## 8. Conclusio

Vor dem Hintergrund des Spannungsfeldes Wissenschaftlichkeit vs. Durchführbarkeit bei Programm Evaluationen, war das Ziel der vorliegenden Arbeit, leicht umsetzbare methodische Maßnahmen zur Qualitätssicherung in diesem Bereich aufzuzeigen. Konkret wurden deshalb bei der Evaluation des Förderprogramms Vienna E-Lecturing zwei Maßnahmen angewendet, deren Nutzen in der vorliegenden Arbeit beschrieben wurde. Im Folgenden werden die wichtigsten Erkenntnisse zusammengefasst sowie weitere Forschungsfragen aufgeworfen.

Im ersten Schritt wurde ein neu entwickeltes matching Verfahren vorgestellt: das Euclidean Distance Matching (EuM). Wenn ein experimentelles Design nicht realisierbar ist, bieten Matching Verfahren i.A. eine gängige Alternative, um trotzdem kausale Effekte der Intervention feststellen zu können. Der konkrete Vorteil des neu entwickelten EuM Verfahrens im Vergleich zu bestehenden üblichen Methoden liegt einerseits in der einfachen Anwendung andererseits in der Möglichkeit mehrere matching Variablen einzubinden, was auch meistens eher der komplexen Realität entspricht (Rogers, 2008). So zeigen auch die Ergebnisse der Anwendung des EuM bei der Evaluation von VEL, dass sich die gematchte Vergleichsgruppe von der Versuchsgruppe zu Beginn des Programms in den relevanten Variablen nicht unterscheidet. Diese Matching Alternative gewinnt v.a. für die Programm Evaluation an Bedeutung, da bisher gebräuchliche Methoden Schwächen aufzeigen. So zeigt sich beispielsweise, dass das propensity score matching erstens ein schlechter Schätzer des „wahren impacts“ eines experimentellen Designs ist und zweitens gegenüber einfacheren, nicht-experimentellen Methoden kaum Vorteile bietet (Wilde & Hollister, 2007). Um

---

Evaluatoren die Entscheidung für ein bestimmtes Verfahren zu erleichtern ist ein systematischer Vergleich aller üblicher quasiexperimenteller Methoden zur Kontrolle von Selektionseffekten notwendig (vgl. Stegmann, 2009). Zur besseren Beurteilung ihrer Qualität ist auch der direkte Vergleich mit einem experimentellen Design hilfreich (Wilde & Hollister, 2007). Diese Vergleiche sind für das EuM noch ausständig.

Im zweiten Schritt wurde der Fokus auf die Bedeutung und den Nutzen personenorientierter Analysen für die Programm Evaluation gelegt. Die aus der Entwicklungspsychologie stammende Idee (Block, 1971, Bergmann and Magnusson, 1997, Spiel, 1998, von Eye and Bergmann, 2003) ist nicht neu, bietet aber auf Grund ihres primären Interesses an der Entwicklung von Personen und weniger an der Veränderung aggregierter Variablenwerte auch für die Programm Evaluation einen wesentlichen Zusatznutzen, der bisher kaum beachtet wurde. Wie auch im Rahmen dieser Arbeit aufgezeigt, wirken Maßnahmen selten auf alle teilnehmenden Personen in gleicher Weise und somit können aggregierte Variablenwerte die tatsächliche Wirkung einer Intervention auf die Personen verfälschen. Bezogen auf die Evaluation von VEL konnten bei der Anwendung personenorientierter Analysen im ersten Schritt – in Anlehnung an das sozial-kognitive Motivationsmodell von Dweck und Leggett (1988) – drei unterschiedliche Typen identifiziert werden: (1) die „Stabil negativ orientierten Studierenden“, die im Fachwissen sowie in allen weiteren untersuchten motivationalen Variablen schlechter abschneiden als (2) die „Kompetenzorientierten Studierenden“ und (3) die „Motivational ausgeglichenen Studierenden“, die sich von den ersteren nur durch eine höhere Leistungszielorientierung unterscheiden. Durch die Anwendung des personenorientierten Ansatzes konnte eine differenzierte Wirkung von VEL

---

aufgedeckt werden. Während ein Studierenden-Typ stark von VEL profitiert, zeigt ein anderer kaum Verbesserungen. Um Ressourcen effizienter einzusetzen könnte in weiterer Folge der Ausschluss der „motivational ausgeglichenen Studierenden“ überlegt werden, da die reguläre Variante von VEL (Vergleichsbedingung) für sie in gleicher Weise wirkt.

Die vorliegende Arbeit zeigt auf, wie ergänzende personenorientierte Analysen den Nutzen und die Qualität der Programm Evaluation erhöhen können. Auf Grund der Vorteile dieser ergänzenden Perspektive findet der personenorientierte Ansatz auch außerhalb der Entwicklungspsychologie immer häufiger Anwendung; so beispielsweise auch in der Gemeindepsychologie (Bogat, 2009) oder eben der Programm-Evaluation (Peck, 2005). Um die wissenschaftliche Qualität von Programm Evaluationen zu sichern, müssen bei der Anwendung personenorientierter Analysemethoden ebenso wie bei variablenorientierten Analysen bestimmte Standards beachtet werden. Wie bereits in dieser Arbeit verdeutlicht ist der Bezug zur Theorie bei der Typenfindung zentral. Weitere Studien, in denen systematisch verschiedene Methoden zur Typenfindung in ihrer Anwendung innerhalb einer Programm Evaluation dargestellt werden sind wünschenswert. So gibt es kaum Literatur, welche beispielsweise die Anwendung einer Konfigurations Frequenz Analyse im Kontext einer Programm Evaluation beschreibt. Ebenso kann ein Vergleich verschiedener Methoden Aufschluss darüber geben, ob simplere personenorientierte Ansätze, wie Extremgruppenvergleiche oder Kategorisierungen (von Eye & Bergman, 2003) als Ergänzung zu allgemeinen variablenorientierten Aussagen zu empfehlen sind.

---

Die vorliegende Dissertation liefert einen konkreten Beitrag um die Qualität von Programm Evaluationen zu erhöhen. Die Relevanz dafür ist nicht nur im Hinblick auf die hohen Qualitätsstandards gegeben, denn die steigende Legitimationspflicht gegenüber politischen Entscheidungsträgern und der Öffentlichkeit (DeGEval, 2010) sowie der immer stärkere Ruf nach evidence based practice (Spiel, 2009) bieten den vorgestellten Maßnahmen ein breites Anwendungsfeld.

---

## 9. Literatur

- Bergman, L. R., & Magnusson, D. (1997). A person-oriented approach in research on developmental psychopathology. *Development and Psychopathology*, 9, 291-319.
- Bierman, K. L. (2006). Commentary on the Pitfalls and Pratfalls of Evaluation Research with Intervention and Prevention Programs. *New Directions for Evaluation*, 110, 87-96.
- Birenbaum, M. & Eylath, S. (1994). Who is afraid of statistics? Correlates of statistics anxiety among students of educational sciences. *Educational Research*, 36, 93-98.
- Block, J. (1971). *Lives through time*. Berkeley: Bancroft Books.
- Bogat, G. A. (2009). Is it necessary to discuss person-oriented research in community psychology? *American Journal of Community Psychology*, 43, 22-34.
- Bortz, J. (1999). *Statistik für Sozialwissenschaftler*. Berlin: Springer.
- Bortz, J. & Döring, N. (2002). *Forschungsmethoden und Evaluation für Human- und Sozialwissenschaftler*. Heidelberg: Springer.
- D'Agostino, Jr.R.B. (2005). Propensity score. In. B.S. Everitt, & D.C. Howell (eds.), *Encyclopedia of Statistics in Behavioral Science* (pp. 1617-1619). Chichester, UK: John Wiley.
- Dehejia, R.H., & Wahba, S. (2002). Propensity score-matching methods for nonexperimental causal studies. *The Review of Economics and Statistics*, 84, 151-161.
- Deutsche Gesellschaft für Evaluation – DeGEval (2004). Standards für Evaluation. Köln: DeGEval.
- Deutsche Gesellschaft für Evaluation – DeGEval (2010). Evaluation und Gesellschaft – Positionspapier 02. Zugriff am 14.02.2010:  
<http://www.degeval.de/calimero/tools/proxy.php?id=21778>
- Dweck, C.S. & Leggett, E.L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review*, 95, 256-273.
- Fagan, A. & Mihalic, S. (2003). Strategies for enhancing the adoption of school-based prevention programs – lessons learned from the blueprints for violence prevention



- 
- replications of the life skills training program. *Journal of Community Psychology*, 31, 235-253.
- Gal, I. & Ginsburg, L. (1994). The role of beliefs and attitudes in learning statistics: Towards an assessment framework. *Journal of Statistics Education*, 2, 1-54.
- Greene, J., Benjamin, L. & Goodyear, L. (2001). The merits of mixing methods in Evaluation. *Evaluation*, 7, 25-44.
- Hager, W., Patry, J.-L. & Brezing, H. (2000) *Handbuch Evaluation psychologischer Interventionsmaßnahmen*. Bern: Hans Huber.
- Hager, W. (2000). Zur Wirksamkeit von Interventionsprogrammen: Allgemeine Kriterien der Wirksamkeit von Programmen in einzelnen Untersuchungen. In Hager, W., Patry, J.-L. & Brezing, H. (Hrsg.), *Handbuch Evaluation psychologischer Interventionsmaßnahmen*. Bern: Hans Huber.
- Joint Committee on Standards for Educational Evaluation, JCSEE (1994). *The program evaluation standards 2<sup>nd</sup> edition*. Thousand Oaks, CA: Sage.
- Lapka, D., Wagner, P., Schober, B., Grading, P. & Spiel, C. (2010). Methodenlehre: Alptraum oder Herausforderung für Psychologiestudierende? Eine Typologie auf Basis des sozialkognitiven Motivationsmodells von Dweck. *Psychologie in Erziehung und Unterricht*, 57 (3).
- Lapka, D., Wagner, P., Schober, B., Grading, P. & Spiel, C. (n.d.). Benefits of the person-oriented perspective for program-evaluation – analyzing for differential treatment effects of the Vienna E-lecturing program. Manuskript eingereicht zur Publikation.
- Lucke, J., Donald, M., Dower, J. & Raphael, B. (2001). Considerations in the Design of a Mixed-Method Cluster - Evaluation of a Community Programme for 'At-Risk' Young People. *Evaluation*, 7, 110-131.
- Magnusson, D. (2000). The individual as the organizing principle in psychological inquiry: A holistic approach. In L. R. Bergman, R. B. Cairns, L.-G. Nilsson & L. Nystedt (Eds.), *Developmental science and the holistic approach* (pp. 33 – 47). Mahwah, NJ: Erlbaum.
- Onwuegbuzie, A.J. (2004). Academic procrastination and statistics anxiety. *Assessment and Evaluation in Higher Education*, 29, 3-19.

- 
- Onwuegbuzie, A.J. & Wilson, V.A. (2003). Statistics anxiety: Nature, etiology, antecedents, effects, and treatments - a comprehensive review of the literature. *Teaching in Higher Education*, 8, 195-209.
- Österreichische Gesellschaft für Psychologie – ÖGP (2009). Empfehlungen des Vorstandes der Österreichischen Gesellschaft für Psychologie zu publikationsbasierten Promotionen. Zugriff am 24.04.2010:  
[http://www.oegp.net/downz/Empfehlungen\\_publicationsorientierte\\_Promotionen\\_Oegp.pdf](http://www.oegp.net/downz/Empfehlungen_publicationsorientierte_Promotionen_Oegp.pdf)
- Patton, M. Q. (1996). *Utilization focused Evaluation* (3th ed.). Thousand Oaks, CA: Sage.
- Peck, L. (2005). Using cluster analysis in program evaluation. *Evaluation Review*, 29, 178-196.
- Rogers, P.J. (2008). Using programme theory to evaluate complicated and complex aspects of interventions. *Evaluation*, 14(1): 29-48
- Rossi, P.H., Freeman, H.E. & Lipsey, M.W. (1999). *Evaluation: A systematic approach*. London: Sage Publication.
- Rubin, D. (1977). Assignment to a treatment group on the basis of a covariate. *Journal of Educational Statistics*, 2, 1-26.
- Scanlon, E., & Issroff, K. (2005). Activity Theory and Higher Education: evaluation learning technologies. *Journal for Computer Assisted Learning*, 21, 430-439.
- Shadish, W.R., Jr., Newman, D.L., Scheirer, M.A., & Wye, Ch. (1995). *Guiding Principles for Evaluators* (New Directions for Program Evaluation, No. 66). Jossey-Basse: San Francisco.
- Scriven, M. (1991). *Evaluation thesaurus* (4th ed.). Newbury Park, CA: Sage.
- Stegman, T. (2009). Analyzing an Active Labor Market Program in Germany: A regional approach. *Journal of MultiDisciplinary Evaluation*, (6)11, 56-70.
- Spiel, C. (1998). Four methodological approaches for the study of stability and change in development. *Methods of Psychological Research* - online 3 (2).  
Online: <http://www.ppm.ipn.uni-kiel.de/mpr/issue5/art2/article.html> [12. 11. 2001].
- Spiel, C., Strohmeier, D., Faradji, S., Schober, B., Gradingner, P., Zens, B., Aichinger, A. & Reimann, R. (2004). Selbstreguliertes Lernen durch Vienna E-Lecturing (VEL). Konzept, Umsetzung und Evaluation. In W. Fröhlich & W. Jütte (Hrsg.),

---

*Qualitätsentwicklung in der postgradualen Weiterbildung. Internationale Entwicklungen und Perspektiven* (S. 377-388). Münster: Waxmann.

Spiel, C. (2005). Program evaluation. In C. B. Fischer & R. M. Lerner (Eds.), *Applied Developmental Science: An encyclopedia of Research, Policies, and Programs* (pp. 879-883). Thousand Oaks, CA: Sage.

Spiel, C., Lapka, D., Gradingner, P., Zodlhofer, E. M., Reimann, R., Schober, B., Wagner, P. & von Eye, A. (2008). A euclidean distance-based matching procedure for non-randomized comparison studies. *European Psychologist*, 13 (3), 180-187.

Spiel, C. (2009). Evidence-based practice: A challenge for European developmental psychology. *European Journal of Developmental Psychology*, 6, 11-33

Spiel, C., Gradingner, P. & Lüftenegger, M. (2009). Grundlagen der Evaluationsforschung. In H. Holling H. & B. Schmitz (Hrsg.), *Handbuch der Psychologischen Methoden und Evaluation*. Göttingen: Hogrefe.

Timmins, P. & Miller, C. (2007). Making evaluations realistic: the challenge of complexity. *Support for Learning*, 22, 9-16.

von Eye, A., & Bergman, L. R. (2003). Research strategies in developmental psychopathology: Dimensional identity and the person-oriented approach. *Development and Psychopathology*, 15 pp. 553-580.

von Eye, A., Bogat, G. A., & Rhodes, J.E. (2006). Variable-oriented and person-oriented perspectives of analysis: The example of alcohol consumption in adolescence. *Journal of Adolescence*, 29, 981-1004.

Wagner, P., Schober, B., Reimann, R., Atria, M. & Spiel, C. (2007). Vienna-Electuring: Trainingskonzept zum selbstregulierten Lernen im Studium. In B. Schmitz & M. Landmann (Hrsg.), *Praxisbuch "Selbstregulation erfolgreich fördern"*. Stuttgart: Kohlhammer Verlag.

Westermann, R. (2002). Merkmale und Varianten von Evaluationen: Überblick und Klassifikation. *Zeitschrift für Psychologie*, 210 (1), 4-26.

Wilde, E. & Hollister, R. (2007). How close is close enough? Evaluating propensity score matching using data from a class size reduction experiment. *Journal of Policy Analysis and Management*, 26 (1), 455-477.

Zeidner, M. (1991). Statistics and mathematics anxiety in social science students: Some interesting parallels. *British Journal of Educational Psychology*, 61, 319-328.

---

## **Abstract**

Bei der Durchführung von Programm Evaluationen stehen Evaluatoren oft vor der Herausforderung, trotz z.T. stark einschränkender Rahmenbedingungen wissenschaftliche Standards einzuhalten. Vor dem Hintergrund dieses Spannungsfeldes liefert die vorliegende Arbeit zwei konkrete methodische Ansätze zur Qualitätssicherung von Programm-Evaluationen. Zum einen wird mit dem Euclidean Distance Matching (EuM) ein neues matching Verfahren vorgestellt, das bei nicht durchführbarer Randomisierung, eine Ähnlichkeit zwischen Kontrollgruppe und Versuchsgruppe in Bezug auf mehrere intervallskalierte Variablen sicherstellt. Gleichzeitig plädiert die vorliegende Arbeit für den ergänzenden Einsatz personenorientierter Analysen in Programm Evaluationen, zur Identifikation differentieller Interventionseffekte. Beide Ansätze werden bei der Evaluation des internetgestützten Hochschulprogramms Vienna E-Lecturing eingesetzt und zeigen in Bezug auf die Qualitätssicherung der Evaluation zufriedenstellende Ergebnisse. Über die Generalisierung auf andere Programm Evaluationen wird jeweils in der Diskussion Stellung genommen.

---

## Curriculum Vitae

Mag. Dominik Lapka

### Persönliche Angaben

Geburtsdatum: 31.08.1981

Geburtsort: Wien

Staatsbürgerschaft: Österreich



### Ausbildung

1991 - 1999	Realgymnasium (1120 Wien)
06/1999	Schulabschluss und Reifeprüfung mit ausgezeichnetem Erfolg
10/1999 - 05/2000	Grundwehrdienst bei der 4. Gardekompanie in Wien
10/2000 - 06/2005	Diplomstudium Psychologie an der Universität Wien in Mindeststudienzeit (Gewählte Schwerpunkte: Wirtschaftspsychologie, Sozialpsychologie)
06/2005	Erfolgreicher Abschluss des Magisterstudiums. Thema der Diplomarbeit: „Wie lernen Studierende? Eine explorative Studie zur Erfassung von Lerntypen im Psychologiestudium“
seit 10/2005	Doktoratsstudium der Psychologie an der Universität Wien

### Berufliche Tätigkeit

03/2004 – 06/2004	Praktikum am Institut für Psychologie der Universität Wien
-------------------	------------------------------------------------------------

- 
- |                   |                                                                                                                                                                                           |
|-------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| 10/2004 - 06/2005 | Studienassistent und Tutor am Institut für<br>Wirtschaftspsychologie, Bildungspsychologie und<br>Evaluation an der Universität Wien                                                       |
| 10/2005 – 10/2009 | Wissenschaftlicher Mitarbeiter im<br>Nationalbankfondprojekt „Vienna E-Lecturing“                                                                                                         |
| seit 10/2005      | Lektor an der Universität Wien; Leitung eines<br>zweisemestrigen Proseminars zur Bildungspsychologie<br>im zweiten Studienabschnitt: „Lebenslanges Lernen –<br>lebensbegleitende Bildung“ |
| seit 10/2009      | Research & Development Consultant bei der Firma<br>Mediacom in Wien                                                                                                                       |

### **Posterpräsentationen**

Lapka, D., Wagner, P., Schober, B., Reimann, R., Grading, P. & Spiel, C. (2006, 21.9.).  
*Methodenlehre: Herausforderung oder Alptraum für Psychologiestudierende?*  
Poster auf dem 45. Kongress der Deutschen Gesellschaft für Psychologie in  
Nürnberg, Deutschland.

Lapka, D., Wagner, P., Schober, B., Reimann, R., Grading, P., Atria, M. & Spiel, C.  
(2006, 29.4). *Vienna E-Lecturing (VEL) – ein internetgestützter Trainingsansatz  
zur Förderung von Fachwissen und Lernkompetenz*. Poster auf der 7.  
Wissenschaftlichen Tagung der Österreichischen Gesellschaft für Psychologie  
(ÖGP) in Klagenfurt, Österreich.

### **Vorträge**

Lapka, D., Wagner, P., Reimann, R., Schober, B., Grading, P. & Spiel, C. (2007,  
August). *Self-monitoring as a central element of Vienna E-Lecturing*. Paper  
presented at the symposium "Improvement of learning and self-regulation with  
different methods of self-monitoring" at the XIIth Biennial Conference for

---

Research on Learning and Instruction (EARLI), Budapest, Hungary, August 28-September 1, 2007.

Lapka, D., Schober, B., Wagner, P., Grading, P. & Spiel, C. (2008, July). *Fostering beneficial self-related cognitions as essential components of self-regulation in a challenging learning context*. Invited paper presented at the symposium "Special features of self-regulated learning" at the XXIXth International Congress of Psychology (ICP), Berlin, Germany, July 20-25, 2008.

### **Artikel & Buchbeiträge**

Lapka, D., Wagner, P., Schober, B., Grading, P. & Spiel, C. (2010). Methodenlehre: Alpträum oder Herausforderung für Psychologiestudierende? Eine Typologie auf Basis des sozialkognitiven Motivationsmodells von Dweck. *Psychologie in Erziehung und Unterricht*, 57(3).

Spiel, C., Lapka, D., Grading, P., Zödlhofer, E. M., Reimann, R., Schober, B., Wagner, P. & von Eye, A. (2008). A Euclidean distance-based matching procedure for non-randomized comparison studies. *European Psychologist*, 13(3), 180-187.

Spiel, C., Reimann, R., Wagner, P., Schober, B., Atria, M., Grading, P. & Lapka, D. (2007). Selbstreguliertes Lernen bei Studierenden fördern: Theorie, Praxis und Evaluation einer blended Lehr-Lern-Umgebung. In A. Gastager & H. Schwetz (Hrsg.). *Pädagogisches Handeln als Balancing zwischen Theorie und Praxis. Beiträge zur Wirksamkeitsforschung im pädagogisch-psychologischen Kontext*. Landau: VEP.